

# André-Oort for a Nonholomorphic Modular Function

Haden Spence

February 13, 2017

## Abstract

The modular case of the André-Oort Conjecture is a theorem of André and Pila, having at its heart the well-known modular function  $j$ . I give an overview of two other ‘nonclassical’ classes of modular function, namely the quasimodular (QM) and almost holomorphic modular (AHM) functions. These are perhaps less well-known than  $j$ , but have been studied by various authors including for example Masser, Shimura and Zagier. It turns out to be sufficient to focus on a particular QM function  $\chi$  and its dual AHM function  $\chi^*$ , since these (together with  $j$ ) generate the relevant fields. After discussing some of the properties of these functions, I go on to prove some Ax-Lindemann results about  $\chi$  and  $\chi^*$ . I then combine these with a fairly standard method of  $\mathfrak{o}$ -minimality and point counting to prove the central result of the paper; a natural analogue of the modular André-Oort conjecture for the function  $\chi^*$ .

## 1 Introduction

Let  $\mathcal{H} = \{\tau \in \mathbb{C} : \text{Im } \tau > 0\}$  be the complex upper half plane. We begin with the classical  $j$ -function, mapping from  $\mathcal{H}$  to  $\mathbb{C}$ , which is well-known to be a *modular function*. It is also transcendental, of course, but nonetheless has rich and beautiful arithmetic properties. For any quadratic point  $\tau \in \mathcal{H}$ , the number  $j(\tau)$  is algebraic over  $\mathbb{Q}$ . Such a point  $j(\tau)$  is called a *special point* or *singular modulus*. The singular moduli are precisely the  $j$ -invariants of elliptic curves with complex multiplication. By a classical theorem of Schneider [15], the imaginary quadratic numbers are the only algebraic elements of  $\mathcal{H}$  whose image under  $j$  is also algebraic.

The special points of  $j$  turn out to be a particular instance of a more general phenomenon. A relation between coordinates in  $\mathcal{H}$  is called a *geodesic relation* if it is of the form  $\tau = c$  for some constant  $c$ , or  $\tau_1 = g\tau_2$  for some  $g \in \text{GL}_2^+(\mathbb{Q})$ . For each  $N \in \mathbb{N}$ , there is a polynomial  $\Phi_N \in \mathbb{Z}[X, Y]$  with the property that

$$\Phi_N(j(\tau), j(g\tau)) = 0,$$

for all  $\tau \in \mathcal{H}$  and any geodesic  $g \in \text{GL}_2^+(\mathbb{Q})$  which, when written as a primitive integer matrix, has determinant  $N$ . So we see that geodesic relations between coordinates  $\tau_i \in \mathcal{H}$  induce algebraic relations between their images  $j(\tau_i) \in \mathbb{C}$ . In fact, geodesic relations are the only algebraic relations in  $\mathcal{H}$  that induce algebraic relations on the  $j$ -side. This fact, known as the Ax-Lindemann theorem for  $j$ , was proven by Pila in [11]; we will discuss it further later.

We call subvarieties of  $\mathbb{C}^n$  which arise in this way  *$j$ -special subvarieties*. So a  $j$ -special subvariety of  $\mathbb{C}^n$  is (an irreducible component of) a variety cut out by some equations of the form  $\Phi_N(z_i, z_j) = 0$  and  $z_i = j(\tau_i)$ , for various values of  $N$  and singular moduli  $j(\tau_i)$ . In general, a  *$j$ -special point* is a zero-dimensional  $j$ -special subvariety, that is, an  $n$ -tuple  $(j(\tau_1), \dots, j(\tau_n))$  where every  $\tau_i$  is a quadratic point.

The  $j$ -special subvarieties of  $\mathbb{C}^n$  are rather sparse; given a random variety  $V \subseteq \mathbb{C}^n$ , we would not expect many  $j$ -special subvarieties to be contained within it. Hence the following finiteness result, proven by Pila in [11]. We call the result Modular André-Oort, since it is a special case of the full André-Oort Conjecture, a statement about general Shimura varieties. The full André-Oort conjecture is known under GRH by work of Edixhoven, Klingler, Ullmo and Yafaev (see for instance [3], [4] and [6]), and is known unconditionally for  $\mathcal{A}_g$ , the moduli space of principally polarised abelian varieties of genus  $g$ ; a result of Tsimerman [18]. In turn, André-Oort is a special case of the far-reaching Zilber-Pink conjecture, so Modular André-Oort forms only a small part of a much larger picture. There is a variety of literature on these topics; good starting points include surveys by Pila [10] and Zannier [22].

**Theorem 1.1** (André/Pila, Modular André-Oort). *Let  $V$  be a subvariety of  $\mathbb{C}^n$ . Then  $V$  contains only finitely many maximal  $j$ -special subvarieties.*

The “maximal” is certainly necessary; in general, a positive-dimensional  $j$ -special variety will always contain infinitely many proper  $j$ -special subvarieties.

The purpose of this paper is to prove an analogue of this result in a slightly different setting. We will be discussing what happens when  $j$  is replaced (or supplemented) by certain ‘modular-like’ functions, which, while not modular functions in the strict sense, exhibit many of the same properties. There are many different such classes of ‘not-quite-modular’ functions, but we will be focussing on two classes in particular: the *quasimodular functions* and the *almost holomorphic modular functions*. Specifically, we will look at a quasimodular function  $\chi$  and a related almost holomorphic modular function  $\chi^*$ , defined as

$$\chi = 1728 \cdot \frac{E_2 E_4 E_6}{E_4^3 - E_6^2}, \quad \chi^* = 1728 \cdot \frac{E_2^* E_4 E_6}{E_4^3 - E_6^2},$$

where  $E_k$  is the usual  $k$ th Eisenstein series and

$$E_2^*(\tau) = E_2(\tau) - \frac{3}{\pi \operatorname{Im} \tau}.$$

See Section 2 for details about the properties of  $\chi$  and  $\chi^*$ . One crucial fact is the existence of modular polynomials  $\Psi_N \in \mathbb{Q}[X, Y, Z]$ , having the property that

$$\Psi_N(\chi^*(g\tau), j(\tau), \chi^*(\tau)) = 0$$

for suitable  $g \in \operatorname{GL}_2^+(\mathbb{Q})$  (compare with the classical modular polynomials  $\Phi_N$ ). In Section 3 we construct these  $\Psi_N$ , using them and the  $\Phi_N$  to construct what we call “ $(j, \chi^*)$ -special” varieties, directly analogous to the “ $j$ -special” varieties discussed above. The  $(j, \chi^*)$ -special varieties exist only inside even Cartesian powers of  $\mathbb{C}$ ; we consider  $\mathbb{C}^{2n}$  as the Zariski closure of  $\pi(\mathcal{H}^n)$ , where

$$\pi : \mathcal{H}^n \rightarrow \mathbb{C}^{2n}$$

is defined by

$$\pi(\tau_1, \dots, \tau_n) = (j(\tau_1), \chi^*(\tau_1), \dots, j(\tau_n), \chi^*(\tau_n)).$$

The central theorem of this paper is the analogue of 1.1 in this setting:

**Theorem 5.8** (André-Oort for  $(j, \chi^*)$ ). *Let  $V \subseteq \mathbb{C}^{2n}$  be a variety. Then  $V$  contains only finitely many maximal  $(j, \chi^*)$ -special subvarieties.*

The proof is quite similar to that of 1.1, using techniques of o-minimality and point-counting. The following “Ax-Lindemann type” result plays a key role. Loosely, it says that all the algebraic sets  $S \subseteq \mathcal{H}^n$  with  $\pi(S) \subseteq V$  are accounted for by the “weakly  $\mathcal{H}$ -special varieties”. These are defined in Section 3; put simply, they are subvarieties of  $\mathcal{H}^n$  cut out by geodesic relations.

**Corollary 4.13.** *Let  $V$  be an irreducible subvariety of  $\mathbb{C}^{2n}$  and let  $\mathcal{Z} = \pi^{-1}(V) \subseteq \mathcal{H}^n$ . Then  $\mathcal{Z}^{\text{alg}}$  is just the union of the weakly  $\mathcal{H}$ -special subvarieties of  $\mathcal{Z}$ .*

Here,  $\mathcal{Z}^{\text{alg}}$  is defined as the union of all positive-dimensional real semialgebraic subsets of  $\mathcal{Z}$ .

It turns out that the proof of this Ax-Lindemann result is where the majority of the work lies; the argument to go from 4.13 to 5.8, done in Section 5, is mostly quite standard. The proof of 4.13 consists of several technical tricks designed to get around the fact that  $\chi^*$  is not holomorphic. Along the way, it turns out that we need an analogous Ax-Lindemann result for the quasimodular function  $\chi$ , which is interesting in its own right, of course. For this, the holomorphicity of  $\chi$  makes the problem somewhat more tractable. We simply take an Ax-Lindemann result of Pila [12], concerning  $j$  and its derivatives, and strengthen it slightly.

In the presence of an Ax-Lindemann theorem for  $\chi$ , it is reasonable to ask whether there is also an André-Oort result in that setting. In fact it is not even obvious that we can formulate such a result. The function  $\chi$ , unlike  $\chi^*$ , does not take algebraic values at quadratic points, so there is no clear notion of what a “ $(j, \chi)$ -special” variety should be.

**Acknowledgements.** Before we proceed, I would like to take a moment to thank the various people who have assisted in the preparation of this document. To my supervisor, Jonathan Pila, an enormous thank you; without Jonathan’s unfailing support and guidance I would be utterly lost. Jonathan has also provided many very helpful suggestions regarding the content and structure of this document specifically. To Alan Lauder, whose suggestion for a miniproject got me interested in nonclassical modular functions in the first place. To my father Derek Spence for proofreading this document several times and making various useful comments and suggestions. To Lincoln College and my college advisor, Ursel Kiehne, who have supported me since my days as an undergraduate. Finally to the EPSRC, who continue to generously fund me throughout my studies.

**Note.** Some of the background material herein also appeared in the author’s Transfer Thesis [17]. It is repeated in this document for the sake of completeness, since the reader is likely not familiar with my transfer thesis!

## 2 Quasimodular and Almost Holomorphic Modular Functions

Let us begin by recalling some basic background about modular functions and Eisenstein series.

**Definition 2.1.** A modular function is a map  $f : \mathcal{H} \rightarrow \mathbb{C}$  with the following properties:

- For any  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$  and any  $\tau \in \mathcal{H}$ , we have  $f(\gamma \cdot \tau) = f(\tau)$ . Here, as usual, elements of the group  $\mathrm{SL}_2(\mathbb{R})$  (more accurately  $\mathrm{PSL}_2(\mathbb{R})$ ) act on  $\mathcal{H}$  via Möbius transformations,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d}.$$

- $f$  is meromorphic on  $\mathcal{H}$ .
- $f$  is “meromorphic at  $\infty$ ”. That is, the Fourier expansion of  $f$ ,

$$f(\tau) = \sum_{k=-\infty}^{\infty} c_k \exp(2k\pi i\tau),$$

has only finitely many negative terms.

One way to construct modular functions is through Eisenstein series. The  $k$ th Eisenstein series  $E_k$  is a function from  $\mathcal{H}$  to  $\mathbb{C}$ , defined as

$$E_k(\tau) = \frac{1}{2} \sum_{\substack{(m,n) \in \mathbb{Z} \\ (m,n)=1}} \frac{1}{(m\tau + n)^k}.$$

For even  $k \geq 4$ , it is easy to see that  $E_k$  converges absolutely, defining a holomorphic function, and further that

$$E_k(\gamma\tau) = (c\tau + d)^k E_k(\tau),$$

where  $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ . (For odd  $k$ , of course, the sum vanishes.) So in particular the function

$$1728 \cdot \frac{E_4^3}{E_4^3 - E_6^2}$$

is invariant under the action of  $\mathrm{SL}_2(\mathbb{Z})$ ; it turns out to be a modular function. In fact this is simply the definition of the  $j$ -function. It is well-known that the denominator

$$\frac{1}{1728}(E_4^3 - E_6^2),$$

which is known as the discriminant function and denoted  $\Delta$ , does not vanish anywhere on  $\mathcal{H}$ , so  $j$  is holomorphic on all of  $\mathcal{H}$ . It turns out that  $j$  is really the only modular function we need to worry about, since the field of modular functions is just  $\mathbb{C}(j)$ .

**Note.** In the remainder of this section we will quite freely use facts proven in Zagier’s excellent paper [21].

So far we have only used the absolutely convergent Eisenstein series  $E_k$ , those with  $k \geq 4$ . Meanwhile, the Eisenstein series  $E_2$  does not converge absolutely, but by taking the terms of the sum in a suitable order, the series does define a holomorphic function  $E_2$ . It does not have the same transformation properties as the other  $E_k$ , but rather satisfies

$$E_2(\gamma\tau) = (c\tau + d)^2 E_2(\tau) - \frac{6i}{\pi} c(c\tau + d).$$

Hence one can see that the modified function

$$E_2^*(\tau) = E_2(\tau) - \frac{3}{\pi \operatorname{Im} \tau}$$

has the usual weight 2 transformation law, that is

$$E_2^*(\gamma\tau) = (c\tau + d)^2 E_2^*(\tau).$$

The functions  $E_2$  and  $E_2^*$  are the prototype examples of, respectively, quasimodular forms and almost holomorphic modular forms.

**Definition 2.2.** A function  $f : \mathcal{H} \rightarrow \mathbb{C}$  is an *almost holomorphic modular form* of weight  $k$  if:

- $f(\tau)$  can be written as a polynomial in  $(\text{Im } \tau)^{-1}$ , with coefficients which are holomorphic functions, bounded as  $\text{Im } \tau \rightarrow \infty$ .
- $f$  satisfies the weight  $k$  transformation law:

$$f(\gamma\tau) = (c\tau + d)^k f(\tau).$$

**Definition 2.3.** A function  $f : \mathcal{H} \rightarrow \mathbb{C}$  is a *quasimodular form* of weight  $k$  if it arises as the constant term (with respect to  $(\text{Im } \tau)^{-1}$ ) of an almost holomorphic modular form of weight  $k$ . Equivalently:

- $f(\tau)$  is a holomorphic function, bounded as  $\text{Im } \tau \rightarrow \infty$ .
- $f$  satisfies the modified transformation law:

$$\frac{f(\gamma\tau)}{(c\tau + d)^k} = f(\tau) + \sum_{k=1}^p f_k(\tau) \left( \frac{c}{c\tau + d} \right)^k,$$

for some holomorphic functions  $f_k$ , bounded as  $\text{Im } \tau \rightarrow \infty$ .

The graded algebra of almost holomorphic modular forms is generated over  $\mathbb{C}$  by  $E_2^*$ ,  $E_4$  and  $E_6$ . The graded algebra of quasimodular forms, similarly, is generated by  $E_2$ ,  $E_4$  and  $E_6$ . In fact, these two graded algebras are isomorphic to each other via the map sending  $E_2^*$  to  $E_2$  and fixing  $E_4$ ,  $E_6$ . One can see this map as that sending an almost holomorphic modular form to its constant coefficient.

For proofs of the various assertions made above, as well as more details about quasimodular and almost holomorphic modular forms in general, see [21]. For this paper, we are more interested in quasimodular and almost holomorphic modular *functions*.

**Definition 2.4.** An almost holomorphic modular (or AHM) function is a quotient of almost holomorphic modular forms of the same weight.

A quasimodular (or QM) function is a quotient of quasimodular forms of the same weight.

The space of AHM functions and the space of QM functions are both fields. We will write  $F^*$  for the field of AHM functions, and  $\tilde{F}$  for the field of QM functions. Each of these fields contains the field of classical modular functions. These have been studied in a few places before, perhaps most notably by Masser in [8]. Masser works with an AHM function he calls  $\psi$ , defined by  $E_2^* E_4 / E_6$ . This function has a singularity at  $i$ , so we work instead with a related function that has no singularities.

Define:

$$f = \frac{E_4 E_6}{\Delta}, \quad \chi = E_2 f, \quad \chi^* = E_2^* f,$$

where  $\Delta$  is again the discriminant function  $(E_4^3 - E_6^2)/1728$ . The function  $f$  is then a meromorphic modular form of weight -2. Since  $\Delta$  does not vanish, none of these three functions have singularities inside of  $\mathcal{H}$ .

Further,  $\chi^*$  is an AHM function and  $\chi$  is a QM function. The function  $\chi$  is holomorphic on  $\mathcal{H}$ , but of course  $\chi^*$  is only real analytic. We note for future use that

$$\chi^* = \chi - \frac{3}{\pi y} f,$$

and (by the transformation properties of  $E_2$ ) that

$$\chi(\gamma\tau) = \chi(\tau) - \frac{6i}{\pi} \frac{c}{c\tau + d} f(\tau),$$

for all  $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{Z})$ .

**Proposition 2.5.** The fields  $F^*$  and  $\tilde{F}$  are characterised by:

$$F^* = \mathbb{C}(j, \chi^*), \quad \tilde{F} = \mathbb{C}(j, \chi).$$

Moreover,  $F^*$  and  $\tilde{F}$  are isomorphic via the map fixing  $j$  and sending  $\chi^*$  to  $\chi$ .

*Proof Sketch.* Zagier proves in [21] that the graded algebras of QM and AHM forms are generated by  $E_4$ ,  $E_6$  and (respectively)  $E_2$  or  $E_2^*$ . Given that fact, it is a simple exercise to write down a generating set for the ‘monomial quotients’ of QM and AHM forms, and see that they are all expressible as rational functions of  $j$  and  $\chi$  or  $\chi^*$ .

The isomorphism of fields is induced directly by the isomorphism between the graded algebras of QM and AHM forms.  $\square$

The following will also be of use.

**Theorem 2.6.** *The functions  $j$ ,  $\chi$  and  $f$  are algebraically independent over  $\mathbb{C}$ .*

*Proof.* Follows easily from the fact that  $E_2$ ,  $E_4$  and  $E_6$  are algebraically independent, which is proven in [21].  $\square$

Our intent is to discuss the special sets corresponding to the functions  $\chi^*$  and  $\chi$ . Such things do exist; they are the subject of the next section.

## 3 Special Sets

### 3.1 New Modular Polynomials

Our discussion of special sets begins with the following proposition involving the construction of some modular polynomials for  $\chi^*$ . Although this follows fairly easily from facts known about  $j$  and its derivatives, together with the upcoming Lemma 3.2, the explicit existence of these polynomials seems not to have been noted before. The construction is very similar to the standard construction of the usual modular polynomials; we follow Zagier [21] closely.

**Proposition 3.1.** *For a positive integer  $N$ , let  $M_N$  be the set of primitive integer matrices  $g \in \mathrm{GL}_2^+(\mathbb{Q})$  with determinant  $N$ . For each such  $N$ , there is a nonzero polynomial  $\Psi_N \in \mathbb{Q}[X, Y, Z]$ , irreducible over  $\mathbb{C}$ , such that*

$$\Psi_N(\chi^*(g\tau), j(\tau), \chi^*(\tau)) = 0$$

for each  $g \in M_N$  and all  $\tau \in \mathcal{H}$ .

*Proof.* The set

$$D_N = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, b, d \in \mathbb{N}, ad = N, 0 \leq b < d, \gcd(a, b, d) = 1 \right\}$$

is a full set of representatives for  $M_N$  under the action of  $\mathrm{SL}_2(\mathbb{Z})$ . That is, for all  $g \in M_N$  there is some  $g' \in D_N$  and  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$  such that  $\gamma g' = g$  (see for instance Lang [7]).

We will consider a polynomial in  $X$ , defined by

$$\prod_{g \in D_N} (X - \chi^*(g\tau)). \quad (1)$$

Clearly (for each  $\tau$ ) this is 0 if and only if  $X$  is  $\chi^*(h\tau)$ , for some  $h \in D_N$ . Thanks to the invariance of  $\chi^*$  under  $\mathrm{SL}_2(\mathbb{Z})$ , this holds if and only if  $X$  is  $\chi^*(h\tau)$  for some  $h \in M_N$ .

Let  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ . For each  $g \in D_N$ , we have  $g \circ \gamma = \gamma' \circ h$ , for some other  $\gamma' \in \mathrm{SL}_2(\mathbb{Z})$  and some  $h \in D_N$ . So by the invariance of  $\chi^*$ , we have

$$\chi^*(g \circ \gamma\tau) = \chi^*(\gamma' \circ h\tau) = \chi^*(h\tau).$$

Thus, the map  $\tau \mapsto \gamma\tau$ , acting on the  $\chi^*(g\tau)$ , simply permutes the  $g \in D_N$ . Each coefficient of  $X$  in the polynomial (1) is a symmetric polynomial in the functions  $\chi^*(g\tau)$ , so each coefficient must be invariant under  $\tau \mapsto \gamma\tau$ . Further, each coefficient is still a polynomial in  $1/y$ , since

$$\mathrm{Im}(g\tau) = \frac{a \mathrm{Im} \tau}{d},$$

where

$$g = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in D_N.$$

Therefore each coefficient of  $X$  is an element of the field of AHM functions  $F^* = \mathbb{C}(j, \chi^*)$ , so can be written as a quotient of complex polynomials in  $j$  and  $\chi^*$ . In each such rational function, we can replace instances of  $j$  and  $\chi^*$  with variables  $Y$  and  $Z$ . If we do this for each coefficient of  $X$ , we get a polynomial

$$\Psi_N^0(X, Y, Z) \in \mathbb{C}(Y, Z)[X]$$

with  $\Psi_N^0(X, j(\tau), \chi^*(\tau)) = 0$  if and only if  $X = \chi^*(g\tau)$  for some  $g \in M_N$ .

Next, note that

$$\chi^*(x + iy) = \frac{E_2 E_4 E_6}{\Delta} - \frac{3}{\pi y} \cdot \frac{E_4 E_6}{\Delta}.$$

Each of the Eisenstein series and  $\Delta$  has a power series expansion in  $q = e^{2\pi iz}$ , with integer coefficients, so  $\chi^*$  itself is a polynomial in  $3/\pi y$  with coefficients that are Laurent series in  $q$  with integer coefficients. The function  $j$  also has a  $q$ -expansion, which is just an integer Laurent series in  $q$ . We will use this to show that  $\Psi_N^0$  is defined over  $\mathbb{Q}$ .

We have

$$\begin{aligned}\Psi_N^0(X, j(\tau), \chi^*(\tau)) &= \prod_{\substack{ad=m \\ d>0}} \prod_{\substack{0 \leq b < d \\ (a,b,d)=1}} \left( X - \chi^* \left( \frac{a\tau + b}{d} \right) \right) \\ &= \prod_{\substack{ad=m \\ d>0}} \prod_{\substack{0 \leq b < d \\ (a,b,d)=1}} \left( X - \sum a_n \zeta_d^{bn} q^{an/d} + \frac{d}{a} \cdot \frac{3}{\pi y} \sum b_n \zeta_d^{bn} q^{an/d} \right),\end{aligned}$$

where  $\zeta_d = e^{2\pi i/d}$  and  $a_n, b_n \in \mathbb{Z}$ . The inner product is a polynomial in  $3/\pi y$ , with coefficients which are Laurent series in  $q^{1/d}$  with coefficients from  $\mathbb{Z}[\frac{d}{a}, \zeta_d]$ . But it is 1-periodic, so the fractional powers of  $q$  must cancel out. Further, the coefficients in the resulting  $q$ -expansions must be in  $\mathbb{Z}[\frac{d}{a}]$ , since every Galois conjugation  $\zeta_d \mapsto \zeta_d^r$ , where  $r \in (\mathbb{Z}/d\mathbb{Z})^*$ , fixes the inner product; the numbers  $b$  and  $rb$  range over the same set.

So each coefficient  $f_k$  of  $X^k$  in  $\Psi_N^0$  is a polynomial in  $3/\pi y$  with coefficients which are rational Laurent series in  $q$ . Each coefficient is also equal to a quotient of polynomials  $p_k$  and  $q_k$  in  $j$  and  $\chi^*$ , thus

$$p_k(j, \chi^*) = f_k \cdot q_k(j, \chi^*).$$

If we compare the coefficients of  $(3/\pi y)^k$  on each side, we get various equalities between  $q$ -expansions. The coefficients of those  $q$ -expansions are  $\mathbb{Q}$ -linear in the coefficients of  $p_k$  and  $q_k$ . So we get a homogeneous system of  $\mathbb{Q}$ -linear equations holding for the coefficients of  $p_k$  and  $q_k$ . This system certainly has a solution since  $p_k$  and  $q_k$  exist. By basic linear algebra, the solution can be chosen to be rational up to scaling, ie.  $p_k$  and  $q_k$  are in  $\lambda\mathbb{Q}[Y, Z]$ , for some  $\lambda$ . In particular,  $p_k/q_k$  can be rewritten as a quotient of rational polynomials.

Thus  $\Psi_N^0 \in \mathbb{Q}(Y, Z)[X]$ . Finally, note that

$$\Psi_N^0(X, j, \chi^*) = \prod_{g \in D_N} (X - \chi^* \circ g) \in F^*[X]$$

is irreducible over  $F^*$ , since no subproduct can have coefficients that are  $\text{SL}_2(\mathbb{Z})$ -invariant. In particular,  $\Psi_N^0(X, Y, Z)$  is irreducible over  $\mathbb{C}(Y, Z)$  as a polynomial in  $X$ . It is also monic in  $X$ , so if we clear the denominators in  $Y$  and  $Z$  exactly, we get an irreducible polynomial  $\Psi_N \in \mathbb{Q}[X, Y, Z]$  having the required properties.  $\square$

In the above, we have made essential use of the fact that  $M_N$  is represented (up to the action of  $\text{SL}_2(\mathbb{Z})$ ) by the finitely many *upper triangular* matrices in  $D_N$ . Since  $\chi^*$  is  $\text{SL}_2(\mathbb{Z})$ -invariant, it is enough that the relation

$$\Psi_N(\chi^*(g\tau), j(\tau), \chi^*(\tau)) = 0$$

holds for  $g \in D_N$ ; that implies the relation for all of  $M_N$ . This is not the case for the QM function  $\chi$ , which only exhibits nice properties with respect to upper triangular elements of  $\text{GL}_2^+(\mathbb{Q})$ . The best we can do is the following.

**Lemma 3.2.** *Let  $g_1, \dots, g_k$  be upper triangular primitive integer matrices with positive determinant, and consider the two fields:*

$$A = F^*(j \circ g_1, \dots, j \circ g_k, \chi^* \circ g_1, \dots, \chi^* \circ g_k)$$

and

$$B = \tilde{F}(j \circ g_1, \dots, j \circ g_k, \chi \circ g_1, \dots, \chi \circ g_k),$$

*considered as fields of real analytic functions, defined locally. Then  $A$  and  $B$  are isomorphic via the map*

$$\chi^* \mapsto \chi, \quad \chi^* \circ g_i \mapsto \chi \circ g_i,$$

*fixing  $j$  and all of the  $j \circ g_i$ .*

*Proof.* The map is clearly a well-defined bijection. If some  $\chi^* \circ g_i$  and  $j \circ g_i$  satisfy a polynomial equation  $p(\chi^* \circ \mathbf{g}, j \circ \mathbf{g}) = 0$ , then (by comparing growth rates) every coefficient of  $1/y$  on the left hand side must vanish. In particular, the constant term  $p(\chi \circ \mathbf{g}, j \circ \mathbf{g})$  must vanish. That is, the same polynomial equation holds for  $\chi \circ \mathbf{g}$  and  $j \circ \mathbf{g}$ , so the map is indeed an isomorphism.  $\square$

It follows that

$$\Psi_N(\chi(g\tau), j(\tau), \chi(\tau)) = 0,$$

for all *upper-triangular* primitive integer matrices of determinant  $N$ . The relation fails for any matrix which is not upper triangular; simply look at the transformation law satisfied by  $\chi$ .

The existence of the polynomials  $\Psi_N$  allows us to say the same thing about  $\chi^*$  that is true of  $j$ , namely: *geodesic relations between coordinates  $\tau_i \in \mathcal{H}$  induce algebraic relations between their images  $j(\tau_i), \chi^*(\tau_i) \in \mathbb{C}$* . Similarly, we can say of  $\chi$  that *upper triangular geodesic relations* induce algebraic relations. Hence, just as for  $j$ , we can talk about the special varieties of  $\mathbb{C}^n$  corresponding to  $\chi^*$  and  $\chi$ .

### 3.2 Special Subvarieties

We will discuss various related types of special subvariety:

- $\mathcal{H}$ -special and weakly  $\mathcal{H}$ -special subvarieties of  $\mathcal{H}^n$ .
- $j$ -special and weakly  $j$ -special subvarieties of  $\mathbb{C}^n$ .
- $(j, \chi^*)$ -special, weakly  $(j, \chi^*)$ -special and weakly  $(j, \chi)$ -special subvarieties of  $\mathbb{C}^{2n}$ .
- $\chi^*$ -special, weakly  $\chi^*$ -special and weakly  $\chi$ -special subvarieties of  $\mathbb{C}^n$ .

We shall see that the weakly  $\chi^*$ -special and weakly  $\chi$ -special subvarieties turn out to be very similar objects. However, the naive approach towards producing a “truly  $\chi$ -special” subvariety does not appear to work. The same goes for  $(j, \chi)$ -special subvarieties.

#### 3.2.1 $\mathcal{H}$ -special and $j$ -special Varieties

We will start with the  $\mathcal{H}$ -special subvarieties.

**Definition 3.3.** Let  $n \in \mathbb{N}$ .

Let  $S_0 \cup S_1 \cup \dots \cup S_k$  be a partition of  $\{1, \dots, n\}$ , where  $k \geq 0$  and  $S_i \neq \emptyset$  for  $i > 0$ . For each  $s \in S_0$ , choose any point  $q_s \in \mathcal{H}$ . For each  $i > 0$ , let  $s_i$  be the least element of  $S_i$  and for each  $s_i \neq s \in S_i$  choose a geodesic matrix  $g_{i,s} \in \mathrm{GL}_2^+(\mathbb{Q})$ . A *weakly  $\mathcal{H}$ -special* subvariety of  $\mathcal{H}^n$  is a set of the form

$$\{(\tau_1, \dots, \tau_n) \in \mathcal{H}^n : \tau_s = q_s \text{ for } s \in S_0, \tau_s = g_{i,s}\tau_{s_i} \text{ for } s \in S_i, s \neq s_i, i = 1, \dots, k\},$$

for some given data  $S_i, q_s, g_{i,s}$ .

A weakly  $\mathcal{H}$ -special subvariety is  *$\mathcal{H}$ -special* if the constant factors  $q_s$  are imaginary quadratic numbers for all  $s \in S_0$ .

**Remark 3.4.** What we call a “weakly  $\mathcal{H}$ -special variety” is elsewhere in the literature referred to as a “geodesic variety”. We have chosen our terminology differently here so that it meshes as closely as possible with the terminology we use for  $j$ -special varieties and so forth.

This definition may look more complex than it actually is. Put more loosely, a weakly  $\mathcal{H}$ -special variety is simply one defined by some number of equations of the form  $\tau_i = q_i$  or  $\tau_i = g_{ij}\tau_j$ , with  $q_i$  constants and  $g_{ij} \in \mathrm{GL}_2^+(\mathbb{Q})$ . If it happens that the  $q_i$  are quadratic points then the variety is fully  $\mathcal{H}$ -special.

This theme continues for all the types of special variety we will define here; a special variety will be defined by some combination of:

- Equations requiring some coordinate to be constant (perhaps a “special” constant).
- Equations coming from the modular polynomials.

These other types of special variety will all arise as (the Zariski closures of) the images of  $\mathcal{H}$ -special varieties under various maps. The easiest to deal with are the well-known  $j$ -special varieties.

**Definition 3.5.** Let  $n \in \mathbb{N}$  and let  $S_0 \cup S_1 \cup \dots \cup S_k$  be a partition of  $\{1, \dots, n\}$ , where  $k \geq 0$  and  $S_i \neq \emptyset$  for  $i > 0$ . For each  $s \in S_0$ , choose a point  $j_s \in \mathbb{C}$ . For each  $i > 0$ , let  $s_i$  be the least element of  $S_i$  and for each  $s_i \neq s \in S_i$  choose a positive integer  $N_{i,s}$ . A weakly  $j$ -special subvariety of  $\mathbb{C}^n$  is an irreducible component of a subvariety of the form

$$\{(z_1, \dots, z_n) \in \mathbb{C}^n : z_s = j_s \text{ for } s \in S_0, \Phi_{N_{i,s}}(z_{s_i}, z_s) = 0 \text{ for } s \in S_i, s \neq s_i, i = 1, \dots, k\}$$

for some given data  $S_i, j_s, N_{i,s}$ .

A weakly  $j$ -special variety is  *$j$ -special* if all of the constant factors  $j_s$  are singular moduli, ie.  $j_s = j(\tau_s)$  for some quadratic  $\tau_s \in \mathcal{H}$ .

The  $j$ -special (resp. weakly  $j$ -special) varieties of  $\mathbb{C}^n$  are precisely those varieties that arise as the image (under the map  $(\tau_1, \dots, \tau_n) \mapsto (j(\tau_1), \dots, j(\tau_n))$ ) of an  $\mathcal{H}$ -special (resp. weakly  $\mathcal{H}$ -special) subvariety of  $\mathcal{H}^n$ . These are the special varieties discussed in 1.1.

### 3.2.2 $(j, \chi^*)$ -special Varieties

The  $(j, \chi^*)$ -special subvarieties are slightly more intricate. We start in the simplest positive dimensional case, considering the variety

$$V'_N = \{(W, X, Y, Z) \in \mathbb{C}^4 : \Phi_N(W, Y) = 0, \Psi_N(X, Y, Z) = 0, \Psi_N(Z, W, X) = 0\} \subseteq \mathbb{C}^4.$$

By counting conditions,  $\dim_{\mathbb{C}} V'_N$  is at most 2. In fact,  $\dim_{\mathbb{C}} V'_N = 2$ . To see this, note that  $V'_N$  contains the set

$$S_g = \{(j(\tau), \chi^*(\tau), j(g\tau), \chi^*(g\tau)) : \tau \in \mathcal{H}\}$$

for any  $g \in M_N$ . Since  $j$  and  $\chi^*$  are algebraically independent,  $S_g$  cannot be contained in any algebraic curve; hence  $\dim_{\mathbb{C}} V'_N > 1$ .

I believe that the variety  $V'_N$  is always irreducible. I have calculated the first few modular polynomials  $\Psi_N$  to see in each case that  $V'_N$  is irreducible, but so far have been unable to prove that this is the case for all  $N$ . So we must instead make the following slightly awkward modification, which fortunately has no effect whatsoever on the wider picture. By real analytic continuation,  $V'_N$  has an irreducible component containing  $S_g$ . Call this component  $V_N$ ; it is still 2-dimensional. Since it contains  $S_g$ , the variety  $V_N$  in fact contains all the  $S_g$ ,  $g \in M_N$ , since we can replace the  $\tau$  in the definition of  $S_g$  by any  $\gamma\tau$ ,  $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ . This has the effect of replacing  $g$  with some other  $h \in M_N$ , and we can get all  $h \in M_N$  like this. Moreover, by 3.2,  $V_N$  contains

$$S'_g = \{(j(\tau), \chi(\tau), j(g\tau), \chi(g\tau)) : \tau \in \mathcal{H}\}$$

for any upper triangular  $g \in M_N$ . This motivates a definition:

**Definition 3.6.** A geodesic upper triangular (or GUT) variety is a weakly  $\mathcal{H}$ -special variety for which all of the relations  $g_{i,s}$  are upper triangular matrices.

The  $V_N$  will form the building blocks of  $(j, \chi^*)$ -special varieties.

**Definition 3.7.** Let  $n \in \mathbb{N}$  and let  $S_0 \cup S_1 \cup \dots \cup S_k$  be a partition of  $\{1, \dots, n\}$ , where  $k \geq 0$  and  $S_i \neq \emptyset$  for  $i > 0$ . For each  $s \in S_0$ , choose  $\tau_s \in \mathcal{H}$  and let  $(j_s, c_s) = (j(\tau_s), \chi^*(\tau_s)) \in \mathbb{C}^2$ . For each  $i > 0$ , let  $s_i$  be the least element of  $S_i$  and for each  $s_i \neq s \in S_i$  choose a positive integer  $N_{i,s}$ . A weakly  $(j, \chi^*)$ -special subvariety of  $\mathbb{C}^{2n}$  is an irreducible component of a subvariety of the form

$$\begin{aligned} \{(w_1, z_1, \dots, w_n, z_n) \in \mathbb{C}^{2n} : (w_s, z_s) = (j_s, c_s) \text{ for } s \in S_0, \\ (w_s, z_s, w_{s_i}, z_{s_i}) \in V_{N_{i,s}} \text{ for } s \in S_i, s \neq s_i, i = 1, \dots, k\}, \end{aligned}$$

for some given data  $S_i$ ,  $(j_s, c_s)$ ,  $N_{i,s}$ .

A weakly  $(j, \chi^*)$ -special variety is  $(j, \chi^*)$ -special if every constant factor  $(j_s, c_s)$  is of the form  $(j(\tau_s), \chi^*(\tau_s))$  for some quadratic point  $\tau_s \in \mathcal{H}$ .

Every weakly  $(j, \chi^*)$ -special variety arises as the Zariski closure of the image of a weakly  $\mathcal{H}$ -special variety under the map

$$(\tau_1, \dots, \tau_n) \mapsto (j(\tau_1), \chi^*(\tau_1), \dots, j(\tau_n), \chi^*(\tau_n)).$$

One can see from the polynomials  $\Psi_N$  that  $\chi^*(\tau)$  is an algebraic number whenever  $\tau \in \mathcal{H}$  is quadratic; this also follows from the work of Masser [8]. This is why truly  $(j, \chi^*)$ -special varieties are called such; all their constant factors are special algebraic numbers.

The weakly  $(j, \chi)$ -special varieties differ from  $(j, \chi^*)$ -special varieties only in their constant factors. The definition is identical, except that the constant factors  $(j_s, c_s)$  are chosen to be of the form  $(j(\tau), \chi(\tau))$ . The weakly  $(j, \chi)$ -special varieties arise as the Zariski closures of the images of GUT varieties under the map

$$(\tau_1, \dots, \tau_n) \mapsto (j(\tau_1), \chi(\tau_1), \dots, j(\tau_n), \chi(\tau_n)).$$

**Remark 3.8.** The polynomials  $\Psi_N$  only work for  $\chi$  and  $\chi \circ g$  when all of the relevant matrices are upper triangular. As a result, numbers  $\chi(\tau)$  with  $\tau$  quadratic are *not* algebraic in general. Diaz has proven and conjectured various results about these points and their transcendence properties in his paper [2]. So  $\chi$  seems not to have any points that we could reasonably call special points, other than perhaps the  $\mathrm{SL}_2(\mathbb{Z})$ -translates of  $i$  and  $e^{2\pi i/3}$ , where  $\chi$  vanishes.

This is why we have not attempted to define a notion of “ $(j, \chi)$ -special variety”; the naive approach does not seem to produce a correct definition and it is not immediately clear whether a correct such notion exists. This is also why, in 3.6, we have not defined any notion of “ $\mathcal{H}$ -special GUT variety”. We might, for instance, have required all of the constant factors to be quadratic, or to be special in some other way, but this seems unlikely to produce a meaningful outcome since  $\chi$  has no apparent special points.



### 3.2.3 $\chi^*$ -special Varieties

The final special varieties we are interested in here are the  $\chi^*$ -special varieties. The idea here is that, while any individual polynomial  $\Psi_N$  introduces a dependence on a  $j$ -coordinate, multiple different relations induced by  $\Psi_N$  can cancel each other out and introduce a dependency that exists between the  $\chi^* \circ g_i$  alone, not involving  $j$ .

Unfortunately, it seems difficult to isolate the specific polynomials that arise in this process. So the easiest definition to use for  $\chi^*$ -special varieties is the following:

**Definition 3.9.** A (weakly)  $\chi^*$ -special variety is an irreducible component of the Zariski closure of the projection of a (weakly)  $(j, \chi^*)$ -special variety onto the coordinates corresponding to  $\chi^*$ -variables.

Equivalently, a  $\chi^*$ -special (resp. weakly  $\chi^*$ -special) variety is an irreducible component of the Zariski closure of the image of an  $\mathcal{H}$ -special (resp. weakly  $\mathcal{H}$ -special) set under the map

$$(\tau_1, \dots, \tau_n) \mapsto (\chi^*(\tau_1), \dots, \chi^*(\tau_n)).$$

As before, the weakly  $\chi$ -special varieties differ from the weakly  $\chi^*$ -special varieties only in their constant factors, and there is no obvious concept of a fully  $\chi$ -special variety.

In what follows, we prove a few results about the possible shape of weakly  $\chi^*$ -special varieties.

**Proposition 3.10.** *Let  $N \geq 2$  and let  $S = \chi^{-1}\{0\}$ . Then there is an upper triangular  $g \in M_N$  such that  $\chi(gS)$  is infinite.*

*Proof.* For any  $\tau \in \mathcal{H}$  which is  $\mathrm{SL}_2(\mathbb{Z})$ -equivalent to  $i$ , the Eisenstein series  $E_6$  is equal to 0. In particular,  $\mathrm{SL}_2(\mathbb{Z}) \cdot i \subseteq S$ . So we only need to show (for some  $g$ ) that  $\chi(g(\gamma \cdot i))$  takes infinitely many values as  $\gamma$  varies. This is easy to see simply by considering matrices of the form

$$g = \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \in U_N, \quad \gamma_n = \begin{pmatrix} 1 & -1 \\ 1 - nN & nN \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Then we get

$$g \cdot \gamma_n = \begin{pmatrix} N & -1 \\ 1 - nN & n \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & N \end{pmatrix},$$

so using the transformation law for  $\chi$ , we have

$$\chi(g(\gamma_n \cdot \tau)) = \chi(\tau/N) - \frac{6i}{\pi} \frac{1 - nN}{(1 - nN)(\tau/N) + n} f(\tau/N),$$

where  $f = E_4 E_6 / \Delta$ . Setting  $\tau = i$ , the above expression clearly takes infinitely many values as  $n$  varies, provided that  $f(i/N) \neq 0$ , which is clear since the only zeros of  $E_4$  and  $E_6$  are  $\mathrm{SL}_2(\mathbb{Z})$ -equivalent to  $i$  or  $\rho$ .  $\square$

**Corollary 3.11.** *The modular polynomial  $\Psi_N(X, Y, Z)$  is nonconstant in  $Y$  for all  $N \geq 2$ .*

*Proof.* Otherwise  $\Psi_N = \Psi_N(X, Z)$ , and then by 3.10,  $\Psi_N(0, Z)$  has infinitely many solutions, and so is identically zero. Since  $\Psi_N$  is irreducible, this would mean that  $\Psi_N(X, Z) = X$ , which is clearly false.  $\square$

**Remark 3.12.** The modular polynomial  $\Psi_1(X, Y, Z)$  is just  $X - Z$ . By the above it is the only modular polynomial which does not depend on  $Y$ .

**Corollary 3.13.** *Let  $n \geq 1$ . For each  $1 \leq i \leq n$ , let  $g_i$  be an upper triangular primitive integer matrix with positive determinant  $N_i$ . Suppose that not all the  $N_i$  are equal to 1. Then the Zariski closure of the set*

$$S = \{(\chi(\tau), \chi(g_1\tau), \dots, \chi(g_n\tau)) : \tau \in \mathcal{H}\}$$

*has complex dimension exactly 2. Similarly, the Zariski closure of*

$$S^* = \{(\chi^*(\tau), \chi^*(g_1\tau), \dots, \chi^*(g_n\tau)) : \tau \in \mathcal{H}\}$$

*has complex dimension exactly 2.*

*Proof.* Consider the  $(j, \chi^*)$ -special variety  $W \subseteq \mathbb{C}^{2n+2}$ , defined by

$$W = \{(X_1, \dots, X_n, Y_1, \dots, Y_n) : (X_1, Y_1, X_i, Y_i) \in N_i\}.$$

$W$  is a 2-dimensional variety and contains the sets

$$\{(j(\tau), \dots, j(g_n(\tau)), \chi^*(\tau), \dots, \chi^*(g_n\tau))\}$$

and

$$\{(j(\tau), \dots, j(g_n(\tau)), \chi(\tau), \dots, \chi(g_n(\tau)))\}.$$

So the sets  $S$  and  $S^*$  are each contained in the projection of  $W$  onto the  $Y_i$ -coordinates (which correspond to  $\chi^*$ -variables). Since  $\dim W = 2$ , the projection certainly has dimension at most 2. So we need to show that  $S$  is not contained in a curve, which is immediate from 3.10 since not all the  $N_i$  are equal to 1. By 3.2,  $S^*$  cannot be contained in any curve either.  $\square$

**Corollary 3.14.** *The only positive-dimensional, proper weakly  $\chi^*$ -special (or indeed weakly  $\chi$ -special) subvarieties of  $\mathbb{C}^2$  are the diagonal  $X = Y$  and the horizontal and vertical lines.*

*Proof.* Immediate.  $\square$

## 4 Ax-Lindemann Theorems

### 4.1 The Pila-Wilkie Counting Theorem

In the remainder of this document we will use, several times, the theory of o-minimal structures, a branch of model theory. There is a brief introduction to this topic in the Appendix. The most vital theorem for us is the following:

**Theorem 4.1** (Pila-Wilkie Counting Theorem). *Let  $Z \subseteq \mathbb{C}^n$  be a definable set in an o-minimal structure  $(\mathbb{R}, \{+, \cdot, \dots\}, \{<, \dots\})$ .*

*For each  $\epsilon > 0$  and each  $k \in \mathbb{N}$  there is a constant  $c(Z, k, \epsilon)$ , such that for every  $T \in \mathbb{N}$ , we have*

$$\#\left\{(x_1, \dots, x_n) \in \overline{\mathbb{Q}}^n \cap Z \setminus Z^{\text{alg}} : \max_i [\mathbb{Q}(x_i) : \mathbb{Q}] \leq k, \max_i \text{Ht}(x_i) \leq T\right\} \leq c(Z, k, \epsilon) T^\epsilon.$$

This may require some explanation. The set  $Z$  here is supposed to be “definable in an o-minimal structure”. For details of what this means, one can see the Appendix. It is enough to know that there is a certain class of subsets of  $\mathbb{C}^n$  which will be called “definable in the o-minimal structure  $\mathbb{R}_{\text{an}, \text{exp}}$ ,” or just “definable”. Crucially, the graphs of the functions  $j$ ,  $\chi$  and  $\chi^*$ , restricted to any  $\text{SL}_2(\mathbb{Z})$ -translate of the standard fundamental domain

$$\mathbb{D} = \left\{ \tau \in \mathcal{H} : -\frac{1}{2} < \text{Re } \tau \leq \frac{1}{2}, |\tau| > 1 \right\},$$

are all definable sets. It follows that for any variety  $V \subseteq \mathbb{C}^{kn}$ , the restricted preimage

$$\pi^{-1}(V) \cap \mathbb{D}^n$$

is a definable set whenever  $\pi : \mathcal{H}^n \rightarrow \mathbb{C}^{kn}$  is some combination of the maps  $j$ ,  $\chi$  and  $\chi^*$ . We will be applying 4.1 to sets of precisely this form.

Given a definable set  $Z$ , we can consider what is called the “transcendental part of  $Z$ ”,  $Z \setminus Z^{\text{alg}}$ , where  $Z^{\text{alg}}$  is the union of all positive-dimensional real semialgebraic subsets of  $Z$ . Pila-Wilkie tells us that the number of algebraic points in  $Z \setminus Z^{\text{alg}}$ , of degree less than some fixed  $k$  and height at most  $T$ , grows more slowly than any positive power of  $T$ . Hence, if we can prove that a given definable set  $Z$  contains ‘too many’ algebraic points of a given height and degree, then  $Z$  must contain an arc of a real algebraic curve.

If  $Z = \pi^{-1}(V) \cap \mathbb{D}^n$  as above, our next task is to find out which real algebraic curves can exist within such preimages. For this we need what is known as an Ax-Lindemann-type result; such results are the topic of this section.

### 4.2 Ax-Lindemann for $j$

In the classical setting, Pila proved the upcoming result in his paper towards André-Oort, [11]. It is called the Ax-Lindemann theorem for  $j$ . Of great interest in its own right, it is also vital to the proof of André-Oort-type results, via Pila-Wilkie. Before we can state it, we will need the following definition:

**Definition 4.2.** Let  $Z \subseteq \mathcal{H}^n$ . A *complex algebraic component*  $A$  of  $Z$  is a connected component of a set of the form

$$W \cap \mathcal{H}^n,$$

for  $W$  an irreducible subvariety of  $\mathbb{C}^n$ , with the property that  $A \subseteq Z$ .

**Theorem 4.3** (Pila, “Ax-Lindemann for  $j$ ”). *Let  $V \subseteq \mathbb{C}^n$  be a variety. Define a map  $\pi : \mathcal{H}^n \rightarrow \mathbb{C}^n$  by*

$$\pi(\tau_1, \dots, \tau_n) = (j(\tau_1), \dots, j(\tau_n)),$$

*and let  $\mathcal{Z} = \pi^{-1}(V)$ .*

*A maximal complex algebraic component of  $\mathcal{Z}$  is weakly  $\mathcal{H}$ -special.*

As we noted in the introduction, this is loosely saying: “the only complex algebraic relations between coordinates in  $\mathcal{H}$  that induce algebraic relations between their  $j$ -images in  $\mathbb{C}^n$  are the geodesic relations.”

For our purposes, the Ax-Lindemann theorem for  $j$  also tells us the following.

**Corollary 4.4.** *Let  $V$ ,  $\pi$  and  $\mathcal{Z}$  be as in 4.3. Then  $\mathcal{Z}^{\text{alg}}$  is simply the union of all positive-dimensional weakly  $\mathcal{H}$ -special subvarieties of  $\mathcal{Z}$ .*

To go from 4.3 to 4.4 one just uses the holomorphicity of  $j$ . A real semialgebraic arc in  $\mathcal{Z}$  is contained in a complex algebraic component of  $\mathcal{Z}$  by analytic continuation.

### 4.3 Quasimodular Ax-Lindemann

For the QM function  $\chi$ , a good portion of the work on Ax-Lindemann results is already done for us. The upcoming result is due to Pila, in [12]. To state it, we will need a definition.

**Definition 4.5.** Let  $\tau_1, \dots, \tau_n$  be elements of some algebraic function field  $\mathbb{C}(W)$ . Then  $\tau_1, \dots, \tau_n$  are called *geodesically dependent* if either:

- Some relation of the form  $\tau_i = g\tau_j$  holds, with  $g \in \text{GL}_2^+(\mathbb{Q})$ , or
- At least one of the  $\tau_i$  is constant.

Otherwise, the  $\tau_i$  are called *geodesically independent*.

**Theorem 4.6** (Pila, Ax-Lindemann with Derivatives). *Suppose that  $\mathbb{C}(W)$  is an algebraic function field and that*

$$\tau_1, \dots, \tau_n \in \mathbb{C}(W)$$

*take values in  $\mathcal{H}$  at  $P \in W$ , and are geodesically independent. Then the  $3n$  functions*

$$j(\tau_1), \dots, j(\tau_n), \quad j'(\tau_1), \dots, j'(\tau_n), \quad j''(\tau_1), \dots, j''(\tau_n)$$

*(considered as functions on  $W$  locally near  $P$ ) are algebraically independent over  $\mathbb{C}(W)$ .*

For our purposes, we need a slightly stronger formulation of this result.

**Theorem 4.7** (Ax-Lindemann with Derivatives, Stronger Form). *Let  $F$  be an irreducible polynomial in  $3n+1$  variables over  $\mathbb{C}$ . Let  $A \subseteq \mathcal{H}^n$  be a complex algebraic component and let  $G$  be the smallest weakly  $\mathcal{H}$ -special variety containing  $A$ . Suppose that  $G$  is a GUT variety and that*

$$F(\tau_1, j(\tau_1), j'(\tau_1), j''(\tau_1), \dots, j(\tau_n), j'(\tau_n), j''(\tau_n)) = 0$$

*for all  $(\tau_1, \dots, \tau_n) \in A$ . Then in fact this holds for all  $(\tau_1, \dots, \tau_n) \in G$ .*

*Proof.* We will work by induction on  $n$ . The case  $n = 1$  is immediate.

In this situation, Theorem 4.6 immediately yields the following. There are  $1 \leq i, j \leq n$  and  $g \in \text{GL}_2^+(\mathbb{Q})$  such that  $\tau_i = g\tau_j$  on  $A$ . Since this is a symmetric condition, we may assume that  $i \neq 1$ . Then without loss of generality,  $i = n$ .

Since  $G$  is a GUT variety,  $g$  is upper triangular. Hence there are algebraic functions  $\phi_1, \phi_2, \phi_3$  (induced by the modular polynomials and their derivatives) such that:

$$j(\tau_i) = \phi_1(j(\tau_j)). \tag{2}$$

$$j'(\tau_i) = \phi_2(j(\tau_j), j'(\tau_j)). \tag{3}$$

$$j''(\tau_i) = \phi_3(j(\tau_j), j'(\tau_j), j''(\tau_j)). \tag{4}$$

Substituting this into  $F$  yields

$$F(\tau_1, j(\tau_1), \dots, j''(\tau_{n-1}), \phi_1(j(\tau_j)), \phi_2(j(\tau_j), j'(\tau_j)), \phi_3(j(\tau_j), j'(\tau_j), j''(\tau_j))) = 0$$

whenever  $(\tau_1, \dots, \tau_{n-1}, g\tau_j) \in A$ . We can then rewrite this as

$$\sigma(\tau_1, j(\tau_1), j'(\tau_1), j''(\tau_1), \dots, j(\tau_{n-1}), j'(\tau_{n-1}), j''(\tau_{n-1})) = 0,$$

for some algebraic function  $\sigma$ . This will hold for all  $(\tau_1, \dots, \tau_{n-1}) \in A'$ , where  $A'$  is the projection of  $A$  onto the first  $n-1$  coordinates.

It is possible that  $\sigma$  is the zero function. If so, then working backwards we see that  $F$  vanishes whenever (2), (3) and (4) hold. In particular,  $F$  vanishes whenever  $\tau_i = g\tau_j$ . Hence it must vanish on  $G$ , as required.

If  $\sigma \neq 0$ , we have more work to do. There is an irreducible polynomial  $p_\sigma$  such that

$$p_\sigma(\sigma(\mathbf{X}), \mathbf{X}) = 0$$

for all  $\mathbf{X}$ . In particular,

$$p_\sigma(0, \tau_1, j(\tau_1), j'(\tau_1), j''(\tau_1), \dots, j(\tau_{n-1}), j'(\tau_{n-1}), j''(\tau_{n-1})) = 0 \quad (5)$$

for all  $(\tau_1, \dots, \tau_{n-1}) \in A'$ . Note that  $p_\sigma(0, \mathbf{X})$  is not the zero polynomial.

We can now appeal to induction to see that (5) holds for all  $(\tau_1, \dots, \tau_{n-1}) \in G'$ , where  $G'$  is the projection of  $G$  onto its first  $n-1$  coordinates. Putting it in different terms: 0 is a root of

$$p_\sigma(X, \tau_1, j(\tau_1), \dots, j''(\tau_{n-1})) \quad (6)$$

whenever  $(\tau_1, \dots, \tau_{n-1}) \in G'$ . We can choose a point  $\mathbf{p} \in A'$ , a  $G'$ -open neighbourhood  $V$  of  $\mathbf{p}$  and a complex-open neighbourhood  $W$  of 0 such that: for all  $\mathbf{q} \in V$ , the only root of (6) within  $W$  is the root 0. However,  $\sigma(\tau_1, j(\tau_1), \dots, j''(\tau_{n-1}))$  is always a root of (6). So for all  $(\tau_1, \dots, \tau_{n-1}) \in V$ , we must have

$$\sigma(\tau_1, j(\tau_1), \dots, j''(\tau_{n-1})) = 0.$$

By analytic continuation, this holds for all  $(\tau_1, \dots, \tau_{n-1}) \in G'$ . Recalling the definition of  $\sigma$ , we get that

$$F(\tau_1, j(\tau_1), \dots, j''(\tau_{n-1}), \phi_1(j(\tau_j)), \phi_2(j(\tau_j), j'(\tau_j)), \phi_3(j(\tau_j), j'(\tau_j), j''(\tau_j))) = 0$$

whenever  $(\tau_1, \dots, \tau_{n-1}) \in G'$ . Hence

$$F(\tau_1, j(\tau_1), \dots, j''(\tau_{n-1}), j(g\tau_j), j'(g\tau_j), j''(g\tau_j)) = 0$$

for all  $(\tau_1, \dots, \tau_{n-1}) \in G'$ . In other words

$$F(\tau_1, j(\tau_1), \dots, j''(\tau_n)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in G$ , as required.  $\square$

For our purposes, we need a version of this result that discusses  $j$  and  $\chi$ , rather than the derivatives of  $j$ , hence the following corollary.

**Corollary 4.8.** *Let  $F$  be an irreducible polynomial in  $3n+1$  variables over  $\mathbb{C}$ . Let  $A \subseteq \mathcal{H}^n$  be a complex algebraic component and let  $G$  be the smallest weakly  $\mathcal{H}$ -special variety containing  $A$ . Suppose that  $G$  is a GUT variety and that*

$$F(\tau_1, j(\tau_1), \chi(\tau_1), f(\tau_1), \dots, j(\tau_n), \chi(\tau_n), f(\tau_n)) = 0$$

*for all  $(\tau_1, \dots, \tau_n) \in A$ . Then in fact this holds for all  $(\tau_1, \dots, \tau_n) \in G$ . (Recall that  $f$  is the function  $E_4 E_6 / \Delta$ , which arises as the  $1/y$  term of  $\chi^*$  and in the transformation law for  $\chi$ .)*

*Proof.* Follows easily from 4.7, using the fact that  $j, \chi, f \in \mathbb{C}(j, j', j'')$  and that  $j(\tau), \chi(\tau), f(\tau)$  are algebraically independent functions over  $\mathbb{C}(\tau)$ .  $\square$

**Note.** In an earlier version of this document, I incorrectly claimed even stronger versions of the above results. The previously stated versions are wrong, but the above is sufficient for our needs. I fully retract both of those strengthened statements; with quite some embarrassment!

#### 4.4 Almost Holomorphic Ax-Lindemann

In the case of  $\chi^*$ , some very basic results of the form of 4.3 were proven in [17], classifying the complex algebraic components that can exist within preimages of varieties. However, since  $\chi^*$  is not holomorphic, we cannot ‘complexify the parameter’ in the usual way to reach an analogue of the real semialgebraic version, 4.4. While a real algebraic arc in  $\mathcal{H}^n$  is certainly contained in a complex algebraic component of  $\mathcal{H}^n$ , there is no guarantee that this algebraic component remains within the preimage of the given variety  $V$ . Fortunately, it turns out that one can use a few tricks to get around this ‘complexification problem.’

We will throughout this section be discussing a map  $\pi : \mathcal{H}^n \rightarrow \mathbb{C}^{2n}$ , defined by

$$\pi(\tau_1, \dots, \tau_n) = (j(\tau_1), \chi^*(\tau_1), \dots, j(\tau_n), \chi^*(\tau_n)).$$

We state our Ax-Lindemann theorem in terms of this map.

**Theorem 4.9** (AHM Ax-Lindemann). *Let  $S$  be an arc of a real algebraic curve in  $\mathcal{H}^n$  and suppose that  $S \subseteq \pi^{-1}(V)$ , where  $V$  is some irreducible variety in  $\mathbb{C}^{2n}$ . Then  $S$  is contained in a weakly  $\mathcal{H}$ -special variety  $G$  with  $G \subseteq \pi^{-1}(V)$ .*

The proof of this is necessarily rather technical, so for ease of reading we have broken it into various smaller chunks. The plan is as follows. Firstly, we deal with the case in which the imaginary part of every complex coordinate is constant on the arc  $S$ . This is the content of Lemma 4.10 and Corollary 4.11. With this done, we can assume that the imaginary part of at least one coordinate (say  $\tau_1$ ) is nonconstant on  $S$ . Hence we can parametrise  $S$  in terms of the imaginary part of  $\tau_1$ .

Using this parametrisation, we will show that a particular algebraic function  $\phi$  in the variables  $\text{Im } \tau_1, j(\tau_i), \chi(\tau_i), f(\tau_i)$  vanishes on  $S$ . If  $\phi$  takes a very specific shape, we can conclude via 4.8. Otherwise, we will see that

$$\text{Im } \tau_1 = \psi(j(\tau_1), \dots, j(\tau_n), \chi(\tau_1), \dots, \chi(\tau_n), f(\tau_1), \dots, f(\tau_n))$$

on  $S$ , for some algebraic function  $\psi$ . In this situation, Lemma 4.12 shows that  $\text{Im } \tau_1$  must be constant on  $S$  after all, which is a contradiction.

**Lemma 4.10.** *Let  $G \subseteq \mathcal{H}^n$  be a GUT variety, let  $F$  be a polynomial in  $2n$  variables, and let  $c_1, \dots, c_n$  be real constants. Suppose that*

$$F\left(j(\tau_1), \chi(\tau_1) - \frac{3}{\pi c_1} f(\tau_1), \dots, j(\tau_n), \chi(\tau_n) - \frac{3}{\pi c_n} f(\tau_n)\right) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in G$ . Then

$$F(j(\tau_1), \chi^*(\tau_1), \dots, j(\tau_n), \chi^*(\tau_n)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in G$ .

*Proof.* By induction on  $n$  we may assume that no coordinate is constant on  $G$ . So up to relabelling coordinates, we have

$$\begin{aligned} G = \{ & (\tau_1, g_{1,1}\tau_1, \dots, g_{1,k_1}\tau_1, \\ & \tau_2, g_{2,1}\tau_2, \dots, g_{2,k_2}\tau_2, \\ & \dots, \\ & \tau_r, g_{r,1}\tau_r, \dots, g_{r,k_r}\tau_r) : \tau_1, \dots, \tau_r \in \mathcal{H} \}, \end{aligned}$$

for some upper triangular matrices  $g_{i,j}$ . Hence

$$\begin{aligned} F \left[ j(\tau_1), \chi(\tau_1) - \frac{3}{\pi d_1} f(\tau_1), \dots, j(g_{1,k_1}\tau_1), \chi(g_{1,k_1}\tau_1) - \frac{3}{\pi d_{1,k_1}} f(g_{1,k_1}\tau_1), \right. \\ \dots, \\ \left. j(\tau_r), \chi(\tau_r) - \frac{3}{\pi d_r} f(\tau_r), \dots, j(g_{r,k_r}\tau_r), \chi(g_{r,k_r}\tau_r) - \frac{3}{\pi d_{r,k_r}} f(g_{r,k_r}\tau_r) \right] = 0, \end{aligned} \tag{7}$$

for some suitable relabelling  $d_i, d_{i,j}$  of the constants  $c_i$ .

All of the  $g_{i,j}$  are upper triangular matrices in  $\text{GL}_2^+(\mathbb{Q})$ . So let us consider a general upper triangular matrix  $g = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ . Let  $A = \gcd(b, d)$  and  $D = ad/A$ . Let  $k, m$  be integers such that  $mb + kd = A$ . For all integers  $t$ , we have

$$\begin{pmatrix} b/A & -k + tb \\ d/A & m + td \end{pmatrix} \cdot \begin{pmatrix} A & -ma \\ 0 & D \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & tD \end{pmatrix}.$$

The leftmost matrix is an element of  $\mathrm{SL}_2(\mathbb{Z})$ . The matrix  $\begin{pmatrix} A & -ma \\ 0 & D \end{pmatrix}$  has the same determinant as  $g$ ; we shall call this matrix  $h$ . Note (taking  $t = 0$  above) that

$$\begin{pmatrix} b/A & -k \\ d/A & m \end{pmatrix} h \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = g.$$

From these matrix equations and the transformation properties of  $j$ ,  $\chi$  and  $f$ , we can easily see that

$$j \left( g \begin{pmatrix} 0 & -1 \\ 1 & tD \end{pmatrix} \tau \right) = j(h\tau), \text{ for all } t. \quad \chi \left( g \begin{pmatrix} 0 & -1 \\ 1 & tD \end{pmatrix} \tau \right) \rightarrow \chi(h\tau) \text{ as } t \rightarrow \infty.$$

$$f \left( g \begin{pmatrix} 0 & -1 \\ 1 & tD \end{pmatrix} \tau \right) \rightarrow 0 \text{ as } t \rightarrow \infty.$$

Also,

$$j \left( \begin{pmatrix} 0 & -1 \\ 1 & tD \end{pmatrix} \tau \right) = j(\tau). \quad \chi \left( \begin{pmatrix} 0 & -1 \\ 1 & tD \end{pmatrix} \tau \right) \rightarrow \chi(\tau) \text{ as } t \rightarrow \infty.$$

$$f \left( \begin{pmatrix} 0 & -1 \\ 1 & tD \end{pmatrix} \tau \right) \rightarrow 0 \text{ as } t \rightarrow \infty.$$

In equation (7), we can replace each  $\tau_i$  by  $\begin{pmatrix} 0 & -1 \\ 1 & tD_i \end{pmatrix} \tau_i$  for suitable fixed  $D_i$  and arbitrary  $t$ . Letting  $t$  tend to infinity we see by continuity of  $F$  that

$$\begin{aligned} & F \left[ j(\tau_1), \chi(\tau_1), \dots, j(h_{1,k_1} \tau_1), \chi(h_{1,k_1} \tau_1), \right. \\ & \quad \dots, \\ & \quad \left. j(\tau_r), \chi(\tau_r), \dots, j(h_{r,k_r} \tau_r), \chi(h_{r,k_r} \tau_r) \right] = 0 \end{aligned}$$

for all  $\tau_i \in \mathcal{H}$  and certain upper triangular matrices  $h_{j,k}$ . By 3.2 (the isomorphism between upper triangular extensions of the fields of QM/AHM functions), we therefore have

$$\begin{aligned} & F \left[ j(\tau_1), \chi^*(\tau_1), \dots, j(h_{1,k_1} \tau_1), \chi^*(h_{1,k_1} \tau_1), \right. \\ & \quad \dots, \\ & \quad \left. j(\tau_r), \chi^*(\tau_r), \dots, j(h_{r,k_r} \tau_r), \chi^*(h_{r,k_r} \tau_r) \right] = 0 \end{aligned} \tag{8}$$

The matrices  $h_{j,k}$  each have the same relation to  $g_{j,k}$  as  $h$  does to  $g$  in the calculation above. In particular, there is  $\gamma_{j,k} \in \mathrm{SL}_2(\mathbb{Z})$  such that

$$\gamma_{j,k} h_{j,k} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = g_{j,k}.$$

So we can replace each  $\tau_i$  in (8) by  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tau_i$  and use the modularity of  $j$  and  $\chi^*$  to see that

$$\begin{aligned} & F \left[ j(\tau_1), \chi^*(\tau_1), \dots, j(g_{1,k_1} \tau_1), \chi^*(g_{1,k_1} \tau_1), \right. \\ & \quad \dots, \\ & \quad \left. j(\tau_r), \chi^*(\tau_r), \dots, j(g_{r,k_r} \tau_r), \chi^*(g_{r,k_r} \tau_r) \right] = 0. \end{aligned}$$

This says precisely that  $F(j(\tau_1), \chi^*(\tau_1), \dots, j(\tau_n), \chi^*(\tau_n)) = 0$  for all  $(\tau_1, \dots, \tau_n) \in G$ .  $\square$

**Corollary 4.11.** *Let  $S$  and  $V$  be as in 4.9. Suppose that the imaginary part of every complex coordinate is constant on  $S$ . Then  $S$  is contained in a weakly  $\mathcal{H}$ -special variety  $G$  with  $G \subseteq \pi^{-1}(V)$ .*

*Proof.* By induction on  $n$ , we may assume that no complex coordinate is constant on  $S$ . So consider the smallest weakly  $\mathcal{H}$ -special variety containing  $S$ , which we will call  $G$ . Since no complex coordinate is constant on  $S$ , the same is true of  $G$ . We want to show  $G \subseteq \pi^{-1}(V)$ .

Consider some coordinate  $\tau_i$  on  $S \subseteq G$ . It takes the form  $\tau_i = x_i + ic_i$ . Suppose that on  $G$ , there is some  $\tau_j$ ,  $j \neq i$  which is related to  $\tau_i$  by some matrix  $g$  which fails to be upper triangular. Then on  $S$ , we have  $\tau_j = g(x_i + ic_i)$ . Since  $\tau_i$  is nonconstant on  $S$ ,  $x_i$  must vary, which then forces  $\text{Im } \tau_j$  to vary since  $g$  is not upper triangular. This is a contradiction. So  $G$  is a GUT variety.

Now pick any of the irreducible polynomials  $F$  which define  $V$ . We have

$$F\left(j(\tau_1), \chi(\tau_1) - \frac{3}{\pi c_1} f(\tau_1), \dots, j(\tau_n), \chi(\tau_n) - \frac{3}{\pi c_n} f(\tau_n)\right) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in S$  and for real constants  $c_i = \text{Im } \tau_i$ . By analytic continuation, this relation must hold on a complex algebraic component  $A$  containing  $S$ . Define  $G$  to be the smallest weakly  $\mathcal{H}$ -special variety containing  $A$ . As previously, we may assume that  $G$  is a GUT variety. Hence we can apply 4.8 to see that

$$F\left(j(\tau_1), \chi(\tau_1) - \frac{3}{\pi c_1} f(\tau_1), \dots, j(\tau_n), \chi(\tau_n) - \frac{3}{\pi c_n} f(\tau_n)\right) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in G$ . By Lemma 4.10, we then have

$$F(j(\sigma_1), \chi^*(\sigma_1), \dots, j(\sigma_n), \chi^*(\sigma_n)) = 0$$

for all  $(\sigma_1, \dots, \sigma_n) \in G$ . This holds for all of the defining polynomials of  $V$ , hence  $S \subseteq G \subseteq \pi^{-1}(V)$  as required.  $\square$

**Lemma 4.12.** *Let  $S$  be an arc of a real algebraic curve in  $\mathcal{H}^n$  and let  $\psi$  be an algebraic function in  $3n$  variables. Suppose that*

$$\text{Im } \tau_1 = \psi(j(\tau_1), \dots, j(\tau_n), \chi(\tau_1), \dots, \chi(\tau_n), f(\tau_1), \dots, f(\tau_n))$$

for all  $(\tau_1, \dots, \tau_n) \in S$ .

Let  $G$  be the smallest weakly  $\mathcal{H}$ -special variety containing  $S$ , and suppose that  $G$  is a GUT variety. Then  $\text{Im } \tau_1$  is constant on  $S$ .

*Proof.* We start with a piece of notation.

**Notation:** The tuple  $(j(\tau_1), \dots, j(\tau_n), \chi(\tau_1), \dots, \chi(\tau_n), f(\tau_1), \dots, f(\tau_n))$  will arise often in what follows, so we abbreviate it as  $\tilde{\pi}(\tau_1, \dots, \tau_n)$ . We will also abbreviate  $y = \text{Im } \tau_1$  throughout. So the first hypothesis of the Lemma may be written as

$$y = \psi(\tilde{\pi}(\tau_1, \dots, \tau_n)).$$

Suppose for a contradiction that  $y = \text{Im } \tau_1$  is nonconstant on  $S$ . Then we can parametrise  $S$  in terms of  $y$ , yielding

$$S = \{(x(y) + iy, u_2(y) + iv_2(y), \dots, u_n(y) + iv_n(y)) : y \in U\}$$

for some open set  $U \subseteq \mathbb{R}$  and algebraic functions  $x, u_i, v_i$ , real-valued on  $U$ .

Since  $S$  is an algebraic arc, we also have some polynomials  $a_i$  such that

$$a_i(x(y), y, u_2(y), v_2(y), \dots, u_n(y), v_n(y)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in S$ . Noting that  $\tau_1 = x(y) + iy$ , and replacing instances of  $y$  with  $\psi$ , we get

$$\begin{aligned} & a_i(\tau_1 - i\psi(\tilde{\pi}(\tau_1, \dots, \tau_n)), \\ & \quad \psi(\tilde{\pi}(\tau_1, \dots, \tau_n)), \\ & \quad u_2(\psi(\tilde{\pi}(\tau_1, \dots, \tau_n))), \\ & \quad v_2(\psi(\tilde{\pi}(\tau_1, \dots, \tau_n))), \\ & \quad \dots, \\ & \quad v_n(\psi(\tilde{\pi}(\tau_1, \dots, \tau_n)))) = 0, \end{aligned} \tag{9}$$

for all  $(\tau_1, \dots, \tau_n) \in S$ . We rewrite the left hand side of this equation as an algebraic function

$$\sigma(\tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n)).$$

Then there is an irreducible polynomial  $p_\sigma$  such that

$$p_\sigma(\sigma(T, \mathbf{J}, \mathbf{X}, \mathbf{F}), T, \mathbf{J}, \mathbf{X}, \mathbf{F}) = 0$$

identically. In particular, since  $\sigma$  vanishes on  $S$ , we have

$$p_\sigma(0, \tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in S$ . By analytic continuation, this holds on a complex algebraic component  $A$  containing  $S$ . Now, the weakly special closure of  $A$  is the same as the weakly special closure of  $S$ , namely  $G$ . Since  $G$  is a GUT variety, we may therefore apply 4.8 to see that

$$p_\sigma(0, \tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in G$ .

In other words, 0 is a root of

$$p_\sigma(X, \tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n)) \tag{10}$$

for all  $(\tau_1, \dots, \tau_n) \in G$ . Since

$$p_\sigma(\sigma(T, \mathbf{J}, \mathbf{X}, \mathbf{F}), T, \mathbf{J}, \mathbf{X}, \mathbf{F}) = 0$$

identically, we know that

$$\sigma(\tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n))$$

is also root of (10) for all  $(\tau_1, \dots, \tau_n) \in G$ .

We can pick a point  $a \in S$ , a  $G$ -open neighbourhood  $W$  of  $a$ , and a complex neighbourhood  $U$  of 0, such that: as  $(\tau_1, \dots, \tau_n)$  varies within  $W$ , the only root of (10) within  $U$  is 0 itself. However, as  $(\tau_1, \dots, \tau_n)$  varies in  $W$ , the function  $\sigma(\tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n))$  remains a root of (10). Since  $\sigma$  vanishes on  $S$ , we can get it arbitrarily close to 0 within  $W$ . In particular, we can get  $\sigma(\tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n))$  to lie within  $U$ . Since it is a root of (10), we must have

$$\sigma(\tau_1, \tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in W$ . By analytic continuation, this holds everywhere on  $G$ , which says that (9) holds on  $G$ .

For notational simplicity, let us suppose that the coordinates which are related to  $\tau_1$  in  $G$  are the first  $k$  coordinates, that is:

$$G = \{(\tau_1, g_2\tau_1, \dots, g_k\tau_1) : \tau_1 \in \mathcal{H}\} \times G',$$

for some other GUT variety  $G'$ . So whenever  $(\tau_1, g_2\tau_1, \dots, g_k\tau_1, \tau_{k+1}, \dots, \tau_n) \in G$ , we also have

$$(\tau_1 + t, g_2(\tau_1 + t), \dots, g_k(\tau_1 + t), \tau_{k+1}, \dots, \tau_n) \in G$$

for every  $t \in \mathbb{Z}$ . Since  $G$  is a GUT set, the  $g_i$  are upper triangular, so the numbers  $g_i(\tau_1 + t)$ , up to translation by an integer, take only finitely many values as  $t$  varies. In particular, since  $j$ ,  $\chi$  and  $f$  are periodic, each of the functions

$$j(g_i(\tau_1 + t)), \quad \chi(g_i(\tau_1 + t)), \quad f(g_i(\tau_1 + t))$$

takes only finitely many values as  $t$  varies. Hence

$$\psi(\tilde{\pi}(\tau_1 + t, g_2(\tau_1 + t), \dots, g_k(\tau_1 + t), \tau_{k+1}, \dots, \tau_n))$$

takes only finitely many values as  $t \in \mathbb{Z}$  varies. If we plug this into (9), we see that

$$a_i(\tau_1 + t - ic, c, u_2(c), v_2(c), \dots, v_n(c)) = 0$$

for some constant  $c$  and infinitely many distinct  $t$ . Thus  $a_i$  is independent of its first coordinate. Since this is true of all the  $a_i$  defining  $S$ , the only possibility for  $S$  is that it is the product of a horizontal line in the  $\tau_1$  plane and points in the other coordinates. So  $y$  is constant on  $S$ , which is a contradiction.  $\square$

With all the above lemmas done, we may finally proceed to the body of the proof of 4.9.

**Proof of 4.9.** By induction on  $n$ , we may assume that no complex coordinate is constant on  $S$ . It might be, however, that the imaginary part of one or more coordinates is constant on  $S$ . If  $\text{Im } \tau_i$  is constant on  $S$  for every  $i$ , then we are in the situation of Lemma 4.11 so we conclude immediately. Hence we may assume without loss of generality that  $\text{Im } \tau_1$  is nonconstant on  $S$ .

Next, let  $G$  be the unique smallest weakly  $\mathcal{H}$ -special subvariety of  $\mathcal{H}^n$  containing  $S$ . There is an element  $\gamma \in \text{SL}_2(\mathbb{Z})^n$  such that  $\gamma G$  is a GUT variety. Then  $\gamma S \subseteq \gamma G$  is still a real semialgebraic arc. By the modularity of  $j$  and  $\chi^*$ ,  $\gamma G \subseteq \pi^{-1}(V)$  if and only if  $G \subseteq \pi^{-1}(V)$ . So by working with  $\gamma S$  we may assume without loss of generality that  $G$  is a GUT variety.

We will write  $y = \text{Im } \tau_1$  throughout, and retain the abbreviation

$$\tilde{\pi}(\tau_1, \dots, \tau_n) = (j(\tau_1), \dots, j(\tau_n), \chi(\tau_1), \dots, \chi(\tau_n), f(\tau_1), \dots, f(\tau_n)).$$



Since  $y$  is nonconstant, we can parametrise  $S$  as

$$S = \{(x(y) + iy, u_2(y) + iv_2(y), \dots, u_n(y) + iv_n(y)) : y \in U\},$$

for some open  $U \subseteq \mathbb{R}$  and algebraic functions  $x, u_i, v_i$ , real-valued on  $U$ .

Consider one of the polynomials  $F$  which defines  $V$ . We have

$$F\left(j(\tau_1), \chi(\tau_1) - \frac{3}{\pi y} f(\tau_1), \dots, j(\tau_n), \chi(\tau_n) - \frac{3}{\pi v_n(y)} f(\tau_n)\right) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in S$ . We can rewrite the left hand side of this equation as an algebraic function

$$\phi(y, \tilde{\pi}(\tau_1, \dots, \tau_n)).$$

Since  $\phi$  is an algebraic function, there is an irreducible polynomial  $p_\phi$  with the property that

$$p_\phi(\phi(T, \mathbf{J}, \mathbf{X}, \mathbf{F}), T, \mathbf{J}, \mathbf{X}, \mathbf{F}) = 0$$

for all  $T, \mathbf{J} = (J_1, \dots, J_n), \mathbf{X} = (X_1, \dots, X_n)$  and  $\mathbf{F} = (F_1, \dots, F_n)$ . In particular, we have that

$$p_\phi(0, y, \tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in S$ . So let us define

$$P(T, \mathbf{J}, \mathbf{X}, \mathbf{F}) = p_\phi(0, T, \mathbf{J}, \mathbf{X}, \mathbf{F}).$$

Note that  $P$  is not the zero polynomial, since  $p_\phi$  is irreducible.

We are going to modify  $P$  as follows. Consider each coefficient of  $T^k$  in  $P$  separately. These are polynomials

$$C_k(\mathbf{J}, \mathbf{X}, \mathbf{F}).$$

For each  $k$ , if

$$C_k(\tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for  $(\tau_1, \dots, \tau_n) \in S$ , then remove this coefficient of  $T^k$  from the polynomial  $P$ . Having done this for each coefficient, we have a modified polynomial which we call  $\tilde{P}$ . Note that we still have

$$\tilde{P}(y, \tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for  $(\tau_1, \dots, \tau_n) \in S$ .

It is possible that  $\tilde{P}$  is the zero polynomial. This happens if and only if every coefficient  $C_k$  has the property that

$$C_k(\tilde{\pi}(\tau_1, \dots, \tau_n)) = 0 \tag{11}$$

for  $(\tau_1, \dots, \tau_n) \in S$ . By analytic continuation, (11) holds for  $(\tau_1, \dots, \tau_n) \in A$ , where  $A$  is the smallest complex algebraic component containing  $S$ . Now, the weakly  $\mathcal{H}$ -special closure of  $A$  is the same as the weakly special closure of  $S$ , which is the GUT variety  $G$ . Hence we can apply 4.8, to see that (11) holds for all  $(\tau_1, \dots, \tau_n) \in G$ .

Since this holds for all  $C_k$ , we have

$$\begin{aligned} & p_\phi(0, Y, \tilde{\pi}(\tau_1, \dots, \tau_n)) \\ &= P(Y, \tilde{\pi}(\tau_1, \dots, \tau_n)) \\ &= 0 \end{aligned}$$

for all  $(\tau_1, \dots, \tau_n) \in G$  and all choices of  $Y$ . In other words, 0 is a root of

$$p_\phi(X, Y, \tilde{\pi}(\tau_1, \dots, \tau_n)) \tag{12}$$

for all  $(\tau_1, \dots, \tau_n) \in G$  and all  $Y$ .

Now we proceed exactly as we did in Lemma 4.12. We can certainly pick a point  $a = (a_1, \dots, a_n) \in S$  such that for all  $(\tau_1, \dots, \tau_n)$  in some  $G$ -open neighbourhood  $W$  of  $a$ , the only root of (12), in some complex neighbourhood  $U$  of 0, is 0 itself. However, we know that

$$X = \phi(Y, \tilde{\pi}(\tau_1, \dots, \tau_n))$$

is a root of (12) identically. Fixing  $Y = \text{Im } a_1$ , we see that  $\phi(\text{Im } a_1, \tilde{\pi}(\tau_1, \dots, \tau_n))$  gets arbitrarily close to 0 within  $W$  (it vanishes at  $a$ ). So as  $(\tau_1, \dots, \tau_n)$  varies within  $W$ ,  $\phi$  is a root of (12), and lies inside of  $U$ . The only such root is 0, so we must have

$$\phi(\text{Im } a_1, \tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in W$ . By analytic continuation, this holds for all  $(\tau_1, \dots, \tau_n) \in G$ . Recalling the definition of  $\phi$ , we get

$$F\left(j(\tau_1), \chi(\tau_1) - \frac{3}{\pi \operatorname{Im} a_1} f(\tau_1), \dots, j(\tau_n), \chi(\tau_n) - \frac{3}{\pi v_n(\operatorname{Im} a_1)} f(\tau_n)\right) = 0$$

for all  $(\tau_1, \dots, \tau_n) \in G$ . Hence we are in the situation of Lemma 4.10, so we get

$$F(j(\tau_1), \chi^*(\tau_1), \dots, j(\tau_n), \chi^*(\tau_n)) = 0,$$

for all  $(\tau_1, \dots, \tau_n) \in G$ , as required.

We have now dealt with the case where  $\tilde{P}$  is the zero polynomial. So we suppose that  $\tilde{P} \neq 0$  and look for a contradiction. Since

$$\tilde{P}(y, \tilde{\pi}(\tau_1, \dots, \tau_n)) = 0$$

for  $(\tau_1, \dots, \tau_n) \in S$ , there is an irreducible factor  $Q$  of  $\tilde{P}$  with this same property.

Suppose some coefficient of  $y^k$  in  $Q$  vanishes on  $S$ . Then we repeat the entire process, removing redundant coefficients to get a polynomial  $\tilde{Q}$ . Again, an irreducible component of  $\tilde{Q}$  must vanish on  $S$ . Then we can remove redundant coefficients from this irreducible component, and so on.

We continue repeating this process until it terminates with an irreducible polynomial

$$R(y, \tilde{\pi}(\tau_1, \dots, \tau_n)),$$

which vanishes on  $S$ , with the property that none of the coefficients of  $y^k$  in  $R$  vanish on  $S$ . If  $R$  were the zero polynomial, then working backwards we see that  $\tilde{P}$  should have been the zero polynomial, which we have assumed is not the case. So  $R \neq 0$ . In particular,  $R$  is nonconstant as a polynomial in  $y$ .

Hence, since none of the coefficients of  $y^k$  in  $R$  vanish on  $S$ , we can extract an algebraic function  $\psi$  such that

$$y = \psi(\tilde{\pi}(\tau_1, \dots, \tau_n))$$

for all  $(\tau_1, \dots, \tau_n) \in S$ . By our earlier comment, we know that the smallest  $\mathcal{H}$ -special variety containing  $S$  is  $G$ , a GUT variety. So we are in the situation of Lemma 4.12, hence  $y$  is constant on  $S$ , which is a contradiction.  $\square$

We can reformulate 4.9 into the following slightly cleaner statement.

**Corollary 4.13.** *Let  $V$  be an irreducible subvariety of  $\mathbb{C}^{2n}$  and let  $\mathcal{Z} = \pi^{-1}(V)$ . Then  $\mathcal{Z}^{alg}$  is just the union of the weakly  $\mathcal{H}$ -special subvarieties of  $\mathcal{Z}$ .*

In the next section we use this to prove the central result of the document.

## 5 André-Oort for $\chi^*$

Since there are no obvious  $\chi$ - or  $(j, \chi)$ -special varieties in  $\mathbb{C}^n$ , it is not clear what an André-Oort statement should look like. We can, however, formulate meaningful André-Oort statements for  $\chi^*$ . In this section we state and prove 5.8, which is the main theorem of the document, an André-Oort theorem for  $j$  and  $\chi^*$ . The map  $\pi$  will throughout be defined as before, namely

$$\pi(\tau_1, \dots, \tau_n) = (j(\tau_1), \chi^*(\tau_1), \dots, j(\tau_n), \chi^*(\tau_n)).$$

The proof follows the standard strategy explicated in [11] very closely, and we will borrow ideas freely from there. Readers familiar with the strategy will be aware of the piece that is currently missing. We need some number-theoretic lower bound in order to force  $\pi^{-1}(V)$  to contain many points of a given height. This will force a real algebraic arc to exist in  $\pi^{-1}(V)$ , so that we can apply the results of the previous section. For the case of  $j$ , the lower bound comes from the size of certain Galois orbits, which are known by a result of Siegel to be sufficiently large. Our approach essentially comes down to that same lower bound of Siegel, but first we have to do some work to ensure that the bound still applies to  $\chi^*$ -special points.

**Proposition 5.1** (Masser). *For a quadratic point  $\tau \in \mathcal{H}$ , we have  $\mathbb{Q}(\chi^*(\tau)) \subseteq \mathbb{Q}(j(\tau))$ .*

*Proof.* Masser proves this in the Appendix of [8] for a function he calls  $\psi$ , which is  $E_2^* E_4 / E_6$ . Since  $\chi^*$  lies in  $\mathbb{Q}(\psi, j)$ , the result follows for  $\chi^*$ .  $\square$

A careful look at Masser's proof of the above yields the following stronger result.

**Proposition 5.2.** *Let  $\tau \in \mathcal{H}$  be a quadratic point and consider the algebraic numbers  $j(\tau)$  and  $\chi^*(\tau)$ . Let  $\sigma$  be a Galois conjugation acting on  $\mathbb{Q}(j(\tau)) \supseteq \mathbb{Q}(\chi^*(\tau))$ . Let  $\tau'$  be a quadratic point such that  $j(\tau') = \sigma(j(\tau))$ . Then  $\chi^*(\tau') = \sigma(\chi^*(\tau))$ .*

*Proof.* This comes entirely from close inspection of Masser's work (the appendix in [8]). Let  $d$  be the discriminant of the quadratic number  $\tau$ , and suppose that  $d$  is not equal to  $3k^2$  for some odd  $k$ . Define some rational functions  $\beta_{i,k}^\tau$  such that  $\beta_{i,k}^\tau(j(\tau))$  are the coefficients of the Taylor expansion of  $\Phi_d$  about the point  $(j(\tau), j(\tau))$ . This we can certainly do, and we get

$$\Phi_d(X, Y) = \sum_{(i,k) \neq (0,0)} \beta_{i,k}^\tau(j(\tau))(X - j(\tau))^i(Y - j(\tau))^k.$$

It appears that the  $\beta_{i,k}^\tau$  will be different rational functions depending on which  $\tau$  we have chosen to expand around. We will show that, for the  $\tau$  and  $\tau'$  defined in the hypotheses of the theorem, we get  $\beta_{i,k}^\tau = \beta_{i,k}^{\tau'}$ . Since  $\Phi_d$  has rational coefficients, any Galois conjugation preserves the left hand side of the above. So we get

$$\begin{aligned} \Phi_d(X, Y) &= \sum \sigma(\beta_{i,k}^\tau(j(\tau)))(X - \sigma(j(\tau)))^i(Y - \sigma(j(\tau)))^k \\ &= \sum \beta_{i,k}^\tau(j(\tau'))(X - j(\tau'))^i(Y - j(\tau'))^k. \end{aligned}$$

We also have

$$\Phi_d(X, Y) = \sum_{(i,k) \neq (0,0)} \beta_{i,k}^{\tau'}(j(\tau'))(X - j(\tau'))^i(Y - j(\tau'))^k,$$

so by uniqueness of Taylor coefficients, the rational functions  $\beta_i^\tau$  and  $\beta_j^{\tau'}$  are equal. On pages 118 and 119 of [8],  $\psi(\tau)$  is expressed as a fixed  $\mathbb{Q}$ -rational function  $p$  in the  $\beta_{i,k}^\tau(j(\tau))$  and  $j(\tau)$ . The equality

$$\psi(\tau) = p(j(\tau), \beta_{i,k}^\tau(j(\tau)))$$

holds whenever  $\tau$  has discriminant  $d$  and  $\beta_{i,j}^\tau$  are the Taylor coefficients of  $\Phi_d$  about  $(j(\tau), j(\tau))$ . Since  $\tau'$  and  $\tau$  have the same discriminant (both satisfy  $\Phi_d(j(\rho), j(\rho)) = 0$ ), this equation holds for both  $\tau$  and  $\tau'$ . Since  $\beta_i^\tau = \beta_i^{\tau'}$  we get

$$\begin{aligned} \sigma\psi(\tau) &= p(\sigma(j(\tau)), \beta_{i,k}^\tau(\sigma(j(\tau)))) \\ &= p(j(\tau'), \beta_{i,k}^\tau(j(\tau'))) \\ &= p(j(\tau'), \beta_{i,k}^{\tau'}(j(\tau'))) \text{ since } \beta_{i,k}^\tau = \beta_{i,k}^{\tau'} \\ &= \psi(\tau'). \end{aligned}$$

When  $\tau$  is  $3k^2$  for some odd  $k$ , the exact same argument still goes through, except the rational function  $p$  is replaced by a different (but still fixed and explicit) rational function. Both  $p$  and  $q$  are written out explicitly on pages 118 and 119 of [8], but we will write them here for completeness<sup>1</sup>:

$$\begin{aligned} p(j, \beta_{i,k}) &= \frac{9j(\beta_{2,0} - \beta_{1,1} + \beta_{0,2})}{\beta_{0,1}} + \frac{3(7j - 6912)}{2(j - 1728)}, \\ q(j, \beta_{i,k}) &= \frac{9j(\beta_{4,0} - \beta_{3,1} + \beta_{2,2} - \beta_{1,3} + \beta_{0,4})}{\beta_{0,1}} + \frac{3(7j - 6912)}{2(j - 1728)}. \end{aligned}$$

In either case we get  $\sigma\psi(\tau) = \psi(\tau')$ . Since  $\chi^* = r(j, \psi)$  for a  $\mathbb{Q}$ -rational function  $r$ , we get  $\sigma\chi^*(\tau) = \chi^*(\tau')$  as required.  $\square$

**Corollary 5.3.** *Let  $K$  be a number field. There are positive constants  $c, \delta > 0$  with the following property. Let  $\tau \in \mathbb{D}$  be a quadratic point of discriminant  $D$ . Then there are  $\gg D^\delta$  distinct quadratic points  $\tau' \in \mathbb{D}$ , of height at most  $cD$ , such that  $(j(\tau'), \chi^*(\tau'))$  is a Galois conjugate, over  $K$ , of the point  $(j(\tau), \chi^*(\tau))$ .*

*Proof.* For quadratic points  $\sigma \in \mathbb{D}$ , let  $H(\sigma)$  be the height of  $\sigma$  and  $D(\sigma)$  the discriminant. It is known that the number of distinct Galois conjugates of  $j(\sigma)$  over  $\mathbb{Q}$  is bounded from below by a positive power of  $D(\sigma)$ . This follows from the Siegel lower bound [16] for class numbers of quadratic fields. See Pila [11] for more details.

Since  $[K : \mathbb{Q}]$  is a fixed constant, the number of Galois conjugates of  $j(\tau)$  over  $K$  is therefore  $\gg D^\delta$ . Each Galois conjugate  $\theta_i$  of  $j(\tau)$  over  $K$  yields a distinct  $\tau_i \in \mathbb{D}$ , such that  $\theta_i(j(\tau)) = j(\tau_i)$ . Moreover,  $D(\tau_i) = D(\tau) = D$ .

<sup>1</sup>The reader may note a strange-looking asymmetry in  $p$  and  $q$ , namely the  $\beta_{0,1}$  in the denominator. Why not  $\beta_{1,0}$ ? Masser in fact proves in his work that  $\beta_{0,1} = \beta_{1,0}$ , so really there is no asymmetry.

By work of Pila [11], there is a constant  $c$  such that, for any  $\sigma \in \mathbb{D}$ ,

$$H(\sigma) \leq cD(\sigma).$$

Hence each  $\tau_i$  has  $H(\tau_i) \leq cD$ .

Finally, by 5.2, we have  $(j(\tau_i), \chi^*(\tau_i)) = (\theta_i(j(\tau)), \theta_i(\chi^*(\tau)))$ .  $\square$

Corollary 5.3 gives us exactly the lower bound we need to work with the Pila-Wilkie theorem. Shortly we will use this bound and the Pila-Wilkie theorem 4.1 to prove our main theorem, 5.8. First we have a proposition demonstrating the ideas in the simplest case; it also serves as the base case for an inductive argument we use in 5.8.

**Proposition 5.4** (André-Oort for  $(j, \chi^*)$ , in 2 Dimensions). *Let  $C \subseteq \mathbb{C}^2$  be an irreducible algebraic curve. Then  $C$  contains only finitely many  $(j, \chi^*)$ -special points.*

*Proof.* Suppose that  $C$  contained infinitely many special points. Since special points are algebraic, this tells us that  $C$  can in fact be defined over  $\overline{\mathbb{Q}}$  and thus over a number field  $K$ .

Define a set  $Z \subseteq \mathbb{D}$  by

$$Z = \{\tau \in \mathbb{D} : (j(\tau), \chi^*(\tau)) \in C\}.$$

Then  $Z$  is definable, and by 4.9 it cannot contain an arc of a real algebraic curve. So  $Z^{\text{alg}}$  is empty. We will show that  $Z \setminus Z^{\text{alg}} = Z$  contains ‘many’ (ie. a positive power of  $T$ ) quadratic points of a given height  $T$ , contradicting the Pila-Wilkie theorem.

Since  $C$  contains infinitely many special points, we have infinitely many distinct quadratic points  $\tau \in \mathbb{D}$  with  $(j(\tau), \chi^*(\tau)) \in C$ . In particular, we can find such a  $\tau$  with arbitrarily large discriminant  $D$ . Hence by 5.3, there are  $\gg D^\delta$  quadratic points  $\tau' \in \mathbb{D}$ , of height at most  $cD$ , such that  $(j(\tau'), \chi^*(\tau'))$  is a Galois conjugate of  $(j(\tau), \chi^*(\tau))$  over  $K$ .

Since it is a Galois conjugate of  $(j(\tau), \chi^*(\tau))$ , we know that  $(j(\tau'), \chi^*(\tau')) \in C$ , hence all of the  $\tau'$  lie in  $Z$ . So there are  $\gg D^\delta$  quadratic points (of height at most  $cD$ ) in  $Z = Z \setminus Z^{\text{alg}}$ , which contradicts the Pila-Wilkie Theorem for any  $\epsilon < \delta$ .  $\square$

In more dimensions, the fundamental ideas for dealing with special points by counting Galois conjugates are exactly the same; we have the following. (Compare with Theorem 11.2 of [11].)

**Proposition 5.5.** *Suppose  $V \subseteq \mathbb{C}^{2n}$  is a variety defined over a number field  $K$ . Write  $V^{\text{sp}}$  for the union of all positive-dimensional  $(j, \chi^*)$ -special subvarieties of  $V$ . Suppose that  $V^{\text{sp}}$  is a variety. Then  $V \setminus V^{\text{sp}}$  contains only finitely many  $(j, \chi^*)$ -special points.*

*Proof.* Let  $\mathcal{Z} = \pi^{-1}(V)$  and  $Z = \mathcal{Z} \cap \mathbb{D}^n$ . Then  $Z$  is definable.

The set  $\mathcal{Z}^{\text{alg}}$  consists of  $\mathcal{Z}^{\text{sp}} = \pi^{-1}(V^{\text{sp}})$  as well as possibly some weakly  $\mathcal{H}$ -special varieties; but the weakly  $\mathcal{H}$ -special varieties can contain no quadratic points. Hence, if we denote by  $N(X, T)$  the number of quadratic points in  $X$  up to height  $T$ , we have

$$N(Z \setminus \mathcal{Z}^{\text{sp}}, T) \leq N(Z \setminus Z^{\text{alg}}, T) \ll_\epsilon T^\epsilon$$

for any  $\epsilon > 0$ ; the last bound coming from the Pila-Wilkie Counting Theorem.

Suppose for a contradiction that  $V \setminus V^{\text{sp}}$  contains infinitely many  $(j, \chi^*)$ -special points. Then we can find quadratic points

$$u = (\tau_1, \dots, \tau_n) \in Z \setminus \mathcal{Z}^{\text{sp}}$$

of arbitrarily large discriminant  $D$ . By 5.3, there are  $\gg D^\delta$  quadratic points  $u' \in \mathbb{D}^n$ , with height at most  $cD$ , such that  $\pi(u')$  is a Galois conjugate of  $\pi(u)$  over  $K$ . This gives us  $\gg D^\delta$  quadratic points (of height at most  $cD$ ) in  $Z \setminus \mathcal{Z}^{\text{sp}}$ . Choosing any  $\epsilon < \delta$ , we get a contradiction to the Pila-Wilkie theorem for sufficiently large  $D$ .  $\square$

So we have some control over the special points that can arise in a given variety. The next step is to deal with the positive-dimensional special subvarieties.

**Definition 5.6.** A  $\mathcal{H}$ -special (or  $(j, \chi^*)$ -special, or  $j$ -special, etc.) variety is called *basic* if it has no constant factors. That is, if the set  $S_0$ , from the definition of a special variety, is empty.

Every weakly  $\mathcal{H}$ -special variety  $S$  arises as the product of a basic  $\mathcal{H}$ -special variety  $B$  with some number of constant factors  $q_i$  (if all the  $q_i$  are quadratic points then  $S$  is fully special). When this happens, we say that  $S$  is the *translate* of  $B$  by the factors  $q_i$ . The following lemma tells us which basic special varieties have translates lying in the preimage of a given variety  $V$ .

**Lemma 5.7.** *Let  $V \subseteq \mathbb{C}^{2n}$  be a variety and define  $\mathcal{Z} = \pi^{-1}(V)$ . There is finite collection  $\mathcal{B}$  of basic  $\mathcal{H}$ -special varieties with the property that every maximal, positive-dimensional, weakly  $\mathcal{H}$ -special subvariety of  $\mathcal{Z}$  is a translate of  $\gamma B$ , for some  $B \in \mathcal{B}$  and  $\gamma \in \mathrm{SL}_2(\mathbb{Z})^k$ .*

*Proof.* This is identical to Proposition 10.2 of [11]. In the presence of 4.9, the proof carries over exactly.  $\square$

Finally, we combine 5.5 and 5.7 in an inductive argument to prove our main theorem.

**Theorem 5.8** (André-Oort for  $(j, \chi^*)$ ). *Let  $V \subseteq \mathbb{C}^{2n}$  be a variety. Then  $V$  contains only finitely many maximal  $(j, \chi^*)$ -special subvarieties.*

*Proof.* There is a subvariety  $\tilde{V} \subseteq V$ , defined over  $\overline{\mathbb{Q}}$ , containing all the algebraic points of  $V$ . So we may assume that  $V$  is defined over  $\overline{\mathbb{Q}}$  (and thus over a number field  $K$ ).

We will proceed by induction on  $n$ . The base case is 5.4. The conclusion holds by 5.5 if  $V^{\mathrm{sp}}$  is variety. So it is sufficient to prove that  $V^{\mathrm{sp}}$  is a variety, under the assumption that 5.8 holds for  $m < n$ .

By 5.7, there are finitely many basic  $\mathcal{H}$ -special varieties,  $B \in \mathcal{B}$ , such that every maximal  $\mathcal{H}$ -special subvariety of  $\pi^{-1}(V)$  is a translate of some  $\gamma B$ . A maximal  $(j, \chi^*)$ -special subvariety of  $V$  is the Zariski closure of  $\pi(S)$ , for some maximal  $\mathcal{H}$ -special subvariety  $S \subseteq \pi^{-1}(V)$ . Therefore any maximal  $(j, \chi^*)$ -special subvariety of  $V$  is the translate (by some special points  $(j(\tau_i), \chi^*(\tau_i))$ ) of one of a finite collection  $\mathcal{C}$  of basic  $(j, \chi^*)$ -special varieties. (The twists by elements of  $\mathrm{SL}_2(\mathbb{Z})$  have no effect since  $j$  and  $\chi^*$  are modular.)

So it is enough to show that, given some basic special  $C \in \mathcal{C}$ , there are only finitely many translates of  $C$  which are maximal  $(j, \chi^*)$ -special subvarieties of  $V$ . Such a  $C$  will be a subvariety of  $\mathbb{C}^{2k}$  for some  $k$ .

The possible translates of  $C$  are elements of  $\mathbb{C}^{2(n-k)}$ , namely the set of points<sup>2</sup>

$$V' = \{(j_1, \chi_1, \dots, j_{n-k}, \chi_{n-k}) : \text{the translate of } C \text{ by } (j_1, \chi_1, \dots, j_{n-k}, \chi_{n-k}) \text{ is contained in } V\}.$$

This is an algebraic subvariety of  $\mathbb{C}^{2(n-k)}$ . The translates of  $C$  which yield special subvarieties of  $V$  are the  $(j, \chi^*)$ -special points of  $V'$ . The translates which yield *maximal* special subvarieties are the  $(j, \chi^*)$ -special points of  $V' \setminus (V')^{\mathrm{sp}}$ . By our inductive assumption, there are only finitely many such points. Thus  $V^{\mathrm{sp}}$ , which consists of finitely many translates of the finitely many basic special varieties in  $\mathcal{C}$ , is a variety. So we can conclude by 5.5.  $\square$

**Corollary 5.9** (André-Oort for  $\chi^*$ ). *Let  $V \subseteq \mathbb{C}^n$  be a variety. Then  $V$  contains only finitely many maximal  $\chi^*$ -special subvarieties.*

*Proof.* Consider a variety  $V' \subseteq \mathbb{C}^{2n}$ , defined as

$$V' = \{(J_1, X_1, \dots, J_n, X_n) : (X_1, \dots, X_n) \in V\}.$$

Given a maximal  $\chi^*$ -special subvariety  $S$  of  $V$ , there is a corresponding  $(j, \chi^*)$ -special subvariety  $S' \subseteq V'$ , such that the projection of  $S'$  onto the  $X_i$  coordinates (which correspond to  $\chi^*$ ) is  $S$ . By 5.8, it is enough to show that  $S'$  is a maximal  $(j, \chi^*)$ -special subvariety of  $V'$ .

Indeed, if  $S'$  were contained in a  $(j, \chi^*)$ -special subvariety  $T \subseteq V'$ , with  $\dim T > \dim S'$ , then by the definition of  $(j, \chi^*)$ -special varieties, there must be a condition on a  $\chi^*$ -coordinate which is relaxed in going from  $S'$  to  $T$ . Hence the projection of  $T$  onto the  $X_i$  coordinates would be a  $\chi^*$ -special subvariety of  $V$  strictly containing  $S$ . Contradiction.  $\square$

## References

- [1] E. BOMBIERI AND W. GUBLER. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006.
- [2] G. DIAZ. Transcendance et indépendance algébrique: liens entre les points de vue elliptique et modulaire. *The Ramanujan Journal*, 4(2):157–199, 2000.
- [3] B. EDIXHOVEN. On the André-Oort conjecture for Hilbert modular surfaces. In *Moduli of abelian varieties (Texel Island, 1999)*, volume 195 of *Progr. Math.*, pages 133–155. Birkhäuser, Basel, 2001.
- [4] B. EDIXHOVEN AND A. YAFAEV. Subvarieties of Shimura varieties. *Ann. of Math. (2)*, 157(2):621–645, 2003.

---

<sup>2</sup>We are being slightly lax with our labelling of coordinates here. The constant factors by which we translate our basic varieties can be in any of the pairs of coordinates in  $(\mathbb{C}^2)^n$ . Since there are only finitely many ways to reorder the coordinates, no issues will arise from allowing the translations to take place in any of the coordinates.

- [5] P. HABEGGER AND J. PILA. Some unlikely intersections beyond André-Oort. *Compos. Math.*, 148(1):1–27, 2012.
- [6] B. KLINGLER AND A. YAFAEV. The André-Oort conjecture. *Ann. of Math. (2)*, 180(3):867–925, 2014.
- [7] S. LANG. *Elliptic functions*, volume 112 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1987. With an appendix by J. Tate.
- [8] D. MASSER. *Elliptic functions and transcendence*. Lecture Notes in Mathematics, Vol. 437. Springer-Verlag, Berlin-New York, 1975.
- [9] Y. PETERZIL AND S. STARCHENKO. Uniform definability of the Weierstrass  $\wp$  functions and generalized tori of dimension one. *Selecta Math. (N.S.)*, 10(4):525–550, 2004.
- [10] J. PILA. O-minimality and Diophantine Geometry. In *Proceedings of the ICM 2014*, volume I, pages 547–572.
- [11] J. PILA. O-minimality and the André-Oort conjecture for  $\mathbb{C}^n$ . *Ann. of Math. (2)*, 173(3):1779–1840, 2011.
- [12] J. PILA. Modular Ax-Lindemann-Weierstrass with derivatives. *Notre Dame J. Form. Log.*, 54(3-4):553–565, 2013.
- [13] J. PILA AND J. TSIMERMAN. Ax-Lindemann for  $A_g$ . *Ann. of Math. (2)*, 179(2):659–681, 2014.
- [14] J. PILA AND A. J. WILKIE. The rational points of a definable set. *Duke Math. J.*, 133(3):591–616, 2006.
- [15] T. SCHNEIDER. *Einführung in die transzendenten Zahlen, Grundlehren LXXXI*. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957.
- [16] C. SIEGEL. Über die Klassenzahl quadratischer Zahlkörper. *Acta Arithmetica*, 1(1):83–86, 1935.
- [17] H. SPENCE. Nonclassical modular functions and their special sets. *Transfer Thesis (for transfer to DPhil status) at the University of Oxford*, 2016.
- [18] J. TSIMERMAN. A proof of the andré-oort conjecture for  $\mathcal{A}_g$ .
- [19] L. VAN DEN DRIES. *Tame topology and o-minimal structures*, volume 248 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1998.
- [20] L. VAN DEN DRIES AND C. MILLER. On the real exponential field with restricted analytic functions. *Israel J. Math.*, 85(1-3):19–56, 1994.
- [21] D. ZAGIER. Elliptic modular forms and their applications. In *The 1-2-3 of modular forms*, Universitext, pages 1–103. Springer, Berlin, 2008.
- [22] U. ZANNIER. *Some problems of unlikely intersections in arithmetic and geometry*, volume 181 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2012. With appendixes by David Masser.

## Appendix: O-minimal Structures

The study of o-minimal structures is a topic in model theory. It has been used to great success by Habegger, Pila, Tsimerman, Masser, Zannier and others to work with problems in functional transcendence, diophantine geometry and other areas in number theory. See for instance [5], [13].

Here we will describe the basic ideas of o-minimal structures and the Pila-Wilkie Counting Theorems. For more details on the basic theory of o-minimal structures, the reader can see [19].

A model-theoretic *structure* is a tuple

$$(X, \{f_i\}, \{R_i\}).$$

The object  $X$  is a set, the *underlying set* of the structure. The other entries in the tuple are collections of functions and relations. The  $f_i$  are some distinguished functions from some cartesian power  $X^{n_i}$  of  $X$  to some other cartesian power  $X^{m_i}$ . The  $R_i$  are relations, subsets of some cartesian power  $X^{k_i}$ .

Structures in this sense are a primary object of study in model theory. Many features and properties can be attributed to different structures, but the relevant notion here is that of a *definable set*. Loosely, a definable set is a subset of  $X^n$  (for some  $n$ ) which can be described precisely in a finite number of statements, mentioning only the following:

- A finite number of the distinguished functions  $f_i$ .
- A finite number of the distinguished relations  $R_i$ .
- A finite number of points chosen from  $X$ .

For instance, if one of the  $f_i$  is a function from  $X$  to itself, the graph of  $f_i$  is a definable subset of  $X^2$ , as is the preimage under  $f_i$  of any given point in  $X$ .

When considered as a collection of sets, the definable subsets have many nice properties, being closed under taking projections, fibres or finite boolean combinations. Indeed, if it is more comfortable, one can forget the definition in terms of functions and relations, and simply say that a structure is a collection of distinguished, ‘definable’ subsets of  $X^n$  with some useful closure properties. We say a function from  $X^m$  to  $X^n$  is definable if its graph is a definable subset of  $X^{m+n}$ .

For the rest of the discussion, we assume that one of the relations  $R_i$  in the structure is an order relation. Such a structure is o-minimal (the ‘o’ standing for ‘order’) if and only if every definable subset of  $X$  is simply a finite union of points and intervals.

**Example.** Let us see some examples of ordered structures.

- The structure  $\overline{\mathbb{R}}$  is the set of real numbers, with addition, multiplication and the usual order as the allowed functions and relations. A definable set is simply a semialgebraic set, so  $\overline{\mathbb{R}}$  is o-minimal.
- If we add the real exponential function to the structure  $\overline{\mathbb{R}}$ , we get  $\mathbb{R}_{\text{exp}}$ , which turns out to be o-minimal.
- Some analytic functions such as  $\sin$ , when added to  $\mathbb{R}$ , immediately produce infinite discrete definable sets like  $\{x \in \mathbb{R} : \sin x = 0\}$ . So we cannot use these functions and preserve o-minimality. However, if we restrict the domains of such functions, o-minimality can be preserved. The structure  $\mathbb{R}_{\text{an,exp}}$  is defined as the expansion of  $\mathbb{R}_{\text{exp}}$  obtained by adding all real analytic functions, with their domains restricted to (a small neighbourhood of) a compact set. It is an important result of van den Dries and Miller [20] that the structure  $\mathbb{R}_{\text{an,exp}}$  is o-minimal.

### The Counting Theorems

When a structure is o-minimal, this has far-reaching implications for the structure as a whole. By definition, we have a nice decomposition of definable subsets of  $X$ , but it is less obvious that we can decompose definable subsets of higher cartesian powers. Definable subsets of  $X^n$  break up into finitely many ‘cells’, which have a nice shape and a natural notion of dimension. This is the so-called Cell Decomposition Theorem, a basic theorem of o-minimality.

Further, definable sets may be parametrised by finite numbers of nice, definable functions, with small derivatives; this is the Parametrisation Theorem. There are many of these basic yet powerful results in the theory of o-minimality, but the key result used in many applications is the Pila-Wilkie Counting Theorem. This is a result giving us control over the distribution of the heights of rational points in a definable set.

The Counting Theorem discusses only the ‘transcendental part’ of a definable set; we had better define what we mean by this.

**Definition.** Let  $Z \subseteq \mathbb{C}^n$ . The algebraic part of  $Z$ , denoted  $Z^{\text{alg}}$ , is the union of all positive-dimensional real semialgebraic subsets of  $Z$ . The transcendental part,  $Z^{\text{tr}}$ , is just  $Z \setminus Z^{\text{alg}}$ .

Now we can state the Counting Theorem.

**Theorem** (Pila-Wilkie Counting Theorem, Version 1). *Let  $Z \subset \mathbb{C}^n$  be a definable set in an o-minimal structure  $(\mathbb{R}, \{+, \cdot, \dots\}, \{<, \dots\})$ .*

*For each  $\epsilon > 0$ , there is a constant  $c(Z, \epsilon)$ , such that for every  $T \in \mathbb{N}$ , we have*

$$\# \left\{ (x_1, \dots, x_n) \in Z^{\text{tr}} \cap \mathbb{Q}^n : \max_i H(x_i) \leq T \right\} \leq c(Z, \epsilon) T^\epsilon.$$

Put simply, this just says that the transcendental part of a definable set contains very few rational points below a given height. Here the *height*  $H$  of a rational number is  $H(a/b) = \max(|a|, |b|)$ , with  $a/b$  in lowest terms.

In discussing special points of modular functions, we need a little more power. The relevant points are quadratic points, not rational points, so the above result cannot say anything about them. Fortunately, there is a stronger version of the Counting Theorem, which counts not just rational points on  $Z^{\text{tr}}$ , but algebraic points as well. We need some notion of the height of an algebraic number, so let  $H$  be the absolute multiplicative height, as defined for instance in 1.5.7 of [1]. Then the stronger result is as follows.

**Theorem** (Pila-Wilkie Counting Theorem, Version 2). *Let  $Z \subset \mathbb{C}^n$  be a definable set in an o-minimal structure  $(\mathbb{R}, \{+, \cdot, \dots\}, \{<, \dots\})$ .*

*For each  $\epsilon > 0$  and each  $k \in \mathbb{N}$  there is a constant  $c(Z, k, \epsilon)$ , such that for every  $T \in \mathbb{N}$ , we have*

$$\# \left\{ (x_1, \dots, x_n) \in Z^{\text{tr}} \cap \overline{\mathbb{Q}}^n : \max_i [\mathbb{Q}(x_i) : \mathbb{Q}] \leq k, \max_i H(x_i) \leq T \right\} \leq c(Z, k, \epsilon) T^\epsilon.$$

These theorems (and in fact a still stronger version involving definable families) are proven in [14, Theorems 1.8-1.10]. They are a vital component at the heart of many of the strategies used in applications of o-minimality.

## Definability of the Modular Functions

To use the Counting Theorems for our purposes requires that the functions in question, like  $j$ ,  $\chi^*$  and  $\tilde{\chi}$ , are definable in some o-minimal structure. In this document, we work in  $\mathbb{R}_{\text{an}, \text{exp}}$ ; ‘definable’ will always mean ‘definable in  $\mathbb{R}_{\text{an}, \text{exp}}$ ’. In any structure whose underlying set is  $\mathbb{R}$  (such as  $\mathbb{R}_{\text{an}, \text{exp}}$ ), we identify  $\mathbb{C}$  with  $\mathbb{R}^2$ ; complex multiplication and addition are then definable operations. So it is meaningful to ask whether a map from  $\mathcal{H}$  to  $\mathbb{C}$  is definable. It is central to the original proof of Pila’s Modular André-Oort theorem for  $j$  that the  $j$ -function, restricted to  $\mathbb{D}$ , is definable in this sense.

A result of Peterzil and Starchenko [9] tells us that a suitable restriction of the Weierstrass  $\wp$  function is definable. Then by the theory of elliptic curves, the Eisenstein series  $E_4$  and  $E_6$  are definable when restricted to  $\mathbb{D}$ . Thus in fact all modular forms and functions are definable on  $\mathbb{D}$ , or any finite union of  $\text{SL}_2(\mathbb{Z})$ -translates of  $\mathbb{D}$ .

For the nonclassical setting discussed in this paper, we need to know that  $\chi^*$  and  $\chi$  are also definable. This is easy enough; it is somewhat well-known that

$$\frac{dE_4}{dz} = 2\pi i \cdot \frac{E_2 E_4 - E_6}{3}.$$

Derivatives of definable functions are definable, so this tells us that  $E_2$  is definable when restricted to  $\mathbb{D}$ . From there it is easy to directly define the function  $E_2^*$ . So all polynomials and rational functions in  $E_2$ ,  $E_2^*$ ,  $E_4$  and  $E_6$  are definable on  $\mathbb{D}$ . In particular, all AHM forms, QM forms, AHM functions and QM functions are definable on  $\mathbb{D}$ . The o-minimal definability of elements of  $F^*$  and  $\tilde{F}$  like  $\chi^*$  and  $\chi$  is one of the primary reasons why we originally chose to focus on QM and AHM functions.